*Red de Centros de Investigación
de la Oficina del Economista Jefe
Banco Interamericano de Desarrollo (BID)
Documento de Trabajo R-352*

THE RETURNS TO HEALTH FOR PERUVIAN URBAN ADULTS: DIFFERENTIALS ACROSS GENDERS, THE LIFE-CYCLE AND AND THE WAGE DISTRIBUTION

BY

Edmundo Murrugarra
University of California at Los Angeles (UCLA)

Martín Valdivia
Grupo de Análisis para el Desarrollo (GRADE)
e-mail: jvaldivi@grade.org.pe

*Red de Centros de Investigación
de la Oficina del Economista Jefe
Banco Interamericano de Desarrollo (BID)
Documento de Trabajo R-352*

This paper is part of the research being developed by the authors under the project *Productivity of Household Investments in Health*, jointly funded by the Consorcio de Investigación Económica (CIE) and the IADB Regional Research Network. This version have benefited from comments by T. Paul Schultz, John Strauss, Bill Savedoff and all participants of the IADB workshop that took place in San José, Costa Rica in March, 1998. The authors also acknowledge the valuable research assistance by Tami Aritomi. All remaining errors are clearly our own responsibility.

Abstract

This report shows evidence about determinants of health status for urban adults and their effects on productivity. Accurate estimation of the effect of health on wages is always difficult to obtain due to endogeneity and measurement error of the health indicators that are available in household surveys for developing countries. The health measure used here is the number of days ill, which involves endogeneity and reporting error problems that are controlled for. The use of household sanitary infrastructure and proxies for health prices, measured by the distance to the health center and the average waiting time for attention at the district level, enabled the construction of an instrument variable estimator for the effects of health on wages. The instruments are statistically significant for all urban individuals. Schooling effects on health are positive and strong for urban males, and the positive effect of schooling on health is clearly increasing with age.

The effect of health on wages is positive and robust, especially for urban males. The larger effects of an additional day sick are found among older self-employed males (-4.3%) and those at the bottom of the hourly earnings distribution (-3.8%), and those in the private sector (-1.8%). These results suggest that health has a stronger impact on the wages of those jobs where productivity and health are closely connected, as in the private sector and the self-employed. The inconclusive results among females indicate the need to work in the development of a model that better express the way in which women insert into the labor market.

Inter-American Development Bank
1300 New York Avenue, N.W.
Washington, D.C. 20577

1. Introduction

The empirical connection between investments in education and wages has been widely analyzed, both in developed and developing countries. In contrast, the returns in the labor market to investments in health have been much less studied, although the conceptual connection has long been present in the development literature, for instance, through the idea of efficiency wages (for an overview, see Strauss and Thomas, 1995). The reasons for this situation are based on conceptual difficulties, as well as on the lack of databases that include both health variables and measures of wages or productivity.

On one side, healthier individuals would tend to be more productive either at work or school, and consequently, they will end up getting higher wages in the labor market. But richer individuals will also tend to be healthier, either because they have more resources to spend on health, or because they have more knowledge about the consequences of their consumption choices and behaviors on their own health. Under these circumstances, it is difficult to infer the impact of health-improving policy interventions on individuals' earnings. On the other side, adequate measures of health status are generally scarce in developing countries, especially for adult individuals. The most commonly available indicators are related to self-reported morbidity (days-ill or days-disabled), but they face serious limitations due to self-reporting biases (see Schultz and Tansel, 1997).

In this study, we focus on the estimation of the effects of health, as measured by the days ill reported by adult individuals over the four weeks prior to the day of the survey, on the average hourly income obtained by wage-earners and self-employed individuals in urban Peru. The database we use comes from the 1994 Peruvian Living Standards Measurement Survey (PLSMS). In principle, it is assumed that the estimated Mincer wage equation has an unobservable health status variable as part of the set of explanatory variables, which is approximated by the observable morbidity indicator. Our objective is twofold. First, we analyze the nature of the interaction between the two classical measures of human capital, education and health, in a wage equation. Second, we examine the effects of health on wages using different sub-samples distinguished by gender, age, location in the wage distribution, and type of employer (public/private; large/small) and employment (blue-collar/white-collar; wage-earner/self-employed).

To properly estimate these effects, we control for endogeneity and measurement error in the available health indicator by using a predicted value for health status with a health equation. The health equation is obtained by solving a household model that has health in the utility function, and a particular production function for health. Individual health outcomes are, thus, a function of individual, family and regional characteristics. To identify health in the wage equation, we have sanitary infrastructure and the availability of health services as the variables that directly affect health but not wages.

The regression results of the (negative) health status equation were very robust for males only, but instruments were also found to be robustly significant for females. For males, we find age differences in the effect of schooling on morbidity. Older males tend to get sick more or longer. But, more importantly, the effect of schooling is uncertain for young males, but is increasing with age, reducing the probability of illness or the number of days sick. Access to adequate housing infrastructure has a negative effect on morbidity, an effect that is very similar across areas (urban/rural) and genders. On the other hand, district-average waiting times before receiving attention in local health centers are positively correlated with urban males' morbidity.

The wage equations for each sector were estimated using Lee (1983)'s MNL-OLS extension of the Heckman two-stage procedure. The instruments used for the second step were the household's asset endowments and local labor market characteristics. Regression results show that healthier individuals receive higher wages, even after controlling for education and income effects. Also, returns to education are slightly smaller when health is included in the wage equation, indicating that previous studies might have been overestimating these returns. For males a clear pattern is observed: the effects of bad health on wages are larger for those self-employed, in the private sector and those at the bottom of the wage distribution. The results for females are less conclusive, probable indicating the need for a better model to reflect the way females insert themselves into the labor market.

The paper is organized as follows. In chapter 2, we describe the model used to derive the wage and health equations. In chapter 3, we discuss the nature of the data in the 1994 PLSMS, in terms of the connection between health and wages. In chapter 4, we describe the econometric model used to estimate both equations, and discuss the strategy followed to overcome usual problems in this kind of estimation. Chapter 5 presents the results for the basic model for wage-earners and self-employed, as well as the differences that result from using rather interesting sub-samples. Finally, chapter 6 includes a summary of the findings and some concluding remarks.

2. The conceptual model

The first step is to describe the way health affects individual and household decisions. This analysis is based on a household model with constrained maximization of a joint utility function, following the framework initiated by Becker (1981). It is assumed that a household with *n* members is run by the household head who maximizes a utility function (U), which depends on the consumption, health and leisure of all members[1],

$$U = U(C^i, h^i, l^i) \qquad i = 1,2,\ldots,n \tag{1}$$

where,

$$C^i = (C_1^i, \ldots, C_j^i, \ldots, C_J^i) \qquad i = 1,2,\ldots,n$$

i.e., C^i is a J dimensional vector, with elements corresponding to a commodity group, h^i denotes the health status and l^i denotes the leisure of member i. It is assumed that the utility function is continuous, strictly increasing, strictly quasi-concave and twice-continuously differentiable in all its arguments. Also, it satisfies the Inada condition, i.e., the marginal utility $U_x \rightarrow \infty$ as $x \rightarrow 0$, for $x = c^i, h^i, l^i$, for all i.

The health status of each household member is determined by a general production function, h .

$$h_i = h_i(C^i, Y^i, l^i, Z^i, X^{-i}, Z^{-i}, F, u^i, u^{-i}) \qquad i = 1,2,\ldots,n \tag{2}$$

where Y^i denotes the consumption of health-related inputs by individual i, Z^i denotes the member's observed characteristics, F denotes the access to sanitary and/or medical infrastructure, and u denotes the vector of unobserved characteristics. Also, X^{-i} denotes the consumption, health and leisure of the other members of the family, and Z^{-i}, u^{-i} denote their vectors of observed and unobserved individual characteristics, respectively. The specific variables that appear in the health production change if the i-th member is an adult, a child or an infant. For instance, in a child's health production function, milk consumption and education of the parents are important components in C^i and Z^{-i}, respectively, although they would probably not be important in an adult's health production function. Since adults tend to take care of themselves, it would be only their own education that matters. In the case of adults, the set of unobservable characteristics include health/nutritional status in earlier years, especially as a child.

The household also faces a full income constraint, which is derived from the time and income constraints,

$$\sum_{j=1}^{J} \sum_{i} p_j c_j^i + \sum_{k=J+1}^{K} \sum_{i} p_k Y_k^i + \sum_{i} w l^i = \sum_{i} w T^i + V = S \tag{3}$$

[1]. This is equivalent to assume that household members have identical preferences, that a dictator rules the household, or generally, a unitary household model. Assuming bargaining to explain intra-household allocations complicates the results without providing additional insights for the goals of the paper.

where (P) represents price, (V) is non-labor income, (W) is the wage rate, (T^a) is the total time available of the adult members, and (S) is the full income. Non-labor income (V) includes net profits of any home enterprise, as well as other rents.

The reduced-form health demand function for adults would have the following form[2]:

$$h^{i*} = h\,(P_C\,,\; P_Y\; S,\; F,\; Z^i\,,\; u^i\;) \tag{4}$$

Although the conditional demand functions have usual properties, that is not true for the reduced-form demand equations in (4). The key point is that consumption affects health too, and substitution effects may attenuate some of the direct effects. For instance, as pointed out by Pitt and Rosenzweig (1986), a decrease in the price of health services P_y or an improvement in sanitary or house infrastructure F could generate substitution in consumption patterns that can reinforce or attenuate the positive health effects of such changes. Consequently, nothing conclusive can be said about the effect of prices, or even sanitary infrastructure upon health status, before the econometric estimation.

Finally, following Mincer (1962) or Mincer and Polacheck (1974), the hourly wage equation is defined as a function of the individuals' working experience (A), education (E), health status (h), and regional variables that characterize local labor markets (L), as in (5).

$$w_{ij} = w\left(A_{ij}\,, E_{ij}\,, h_{ij}\,, L_j\,; e_{ij}\right) \tag{5}$$

The specific functional forms for equations (4) and (5) to be used in the empirical analysis are discussed in section 4 below.

3. **The Peruvian data.**

The 1994 PLSMS contains information on adult morbidity and net income for wage-earners and the self-employed, as well as on the characteristics of all individuals in urban and rural households[3]. The total sample size is 19,284 individuals organized in 3,623 households. We restrict our analysis to urban areas because rural labor markets would have a very different structure, and its modeling would probably require a

---

2 . The health production functions are assumed to be twice-continuously differentiable, strictly increasing, and strictly concave function in all arguments. Then, the constraint set formed by (2), and the full income constraint (3) is convex and the optimization of (1) yields a unique solution. Assuming that the health production functions satisfy the Inada condition guarantees the solution to be interior.

3 . See Grosh and Gleewe (1995) for more details on LSMS surveys.

more complicated structure[4]. Further restricting our analysis to adults (16 to 60 years old), our working sample size comes down to 6610 individuals, of which 3,102 are males.

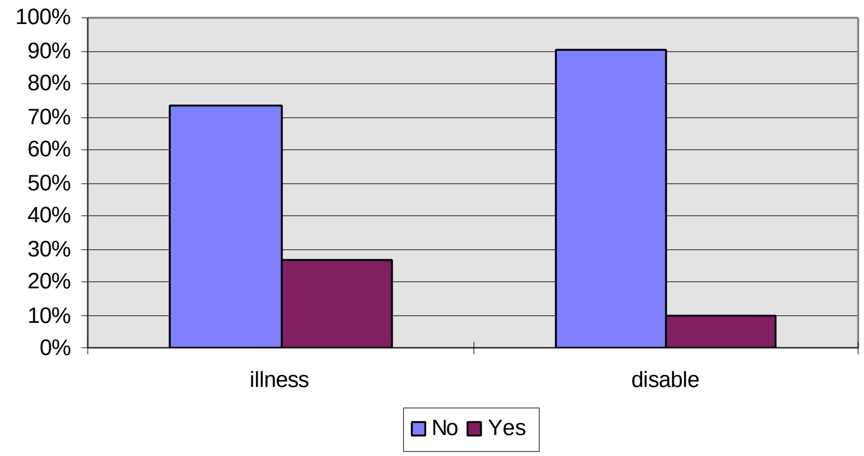
The health measure used in this report is the "Number of days sick or injured in the last four weeks". Typically, LSMS Surveys include a sequence of questions on health and sickness, asking first "Have you had any illness or injury during the past four weeks?," and continuing with the number of days sick, and the number of days disabled due to sickness and other related information. Among these variables, we choose the number of days sick because we presume that it contains more information about the latent health status of individuals than the simple indicator about illness events, that is, we presume that an individual with less days sick in the last four weeks is generally *healthier*.

The typical problem with self-reported health status indicators is that they are contaminated by other individual characteristics such as education and other measures related to the ability to detect illnesses. Between two similar individuals, the one with more unobserved skills to detect an illness would tend to report a larger number of days-ill than the less skilled. Under this type of endogeneity, a positive bias is expected for the coefficients of the health variables. Alternatively, Schultz and Tansel (1997) have worked with the number of days *disabled* for Ghana and Cote d'Ivoire, because it is assumed that such an indicator is less contaminated by this problem, since indicators of functional disability are less subjective. Such a measure, however, loses information on the morbidity of individuals, since those sick but not disabled are considered as healthy as those who experienced no illness at all. In this sense, there is a tradeoff between using a more strict measure that loses information, or using a measure that is more subjective but also more informative.

In the PLSMS, 27% of the individuals aged between 16 and 60 reported having been ill or an accident during the four weeks previous to the survey (see figure 1, below). About a third those individuals report having been disabled as a result of that illness or accident. Despite the potential reporting bias, we choose days-ill in the hourly wage regressions, because it contains valuable information about the latent health status of individuals, especially in the case of the illness events that did not affect functionality.

[4]. Valdivia and Robles (1997) investigated the particular nature of Peruvian labor markets. A relevant feature for the purposes of this study is that most economic units are very small and mainly use family labor. Wage labor markets are used only occasionally or according to the seasonality of agricultural labor demand, to generate some monetary income.

Figure 1: Self-reported ocurrences of illness and disability for Peruvian adults



The variable days-ill is not only censored at zero, but also at 28, since the question refers to the number of days that the individual was ill during the 4 weeks prior to the survey. As mentioned before, 73% of surveyed individuals did not report any sickness. Figure 2 shows the distribution of individuals by the number of days ill, where only those with one or more days are included. Note that, of these, 14% reported being sick for the entire period (28 days). Overall, the average number of days ill is about 2.5 days, but the number is positively correlated with age, as can be seen in Figure 3. For instance, the average number of days ill is below 2 for individuals under 30, but rises above 4 for individuals over 50. This trend is not only the result of a higher probability of getting sick as you get older, but also that older people tend to get sick for longer periods.

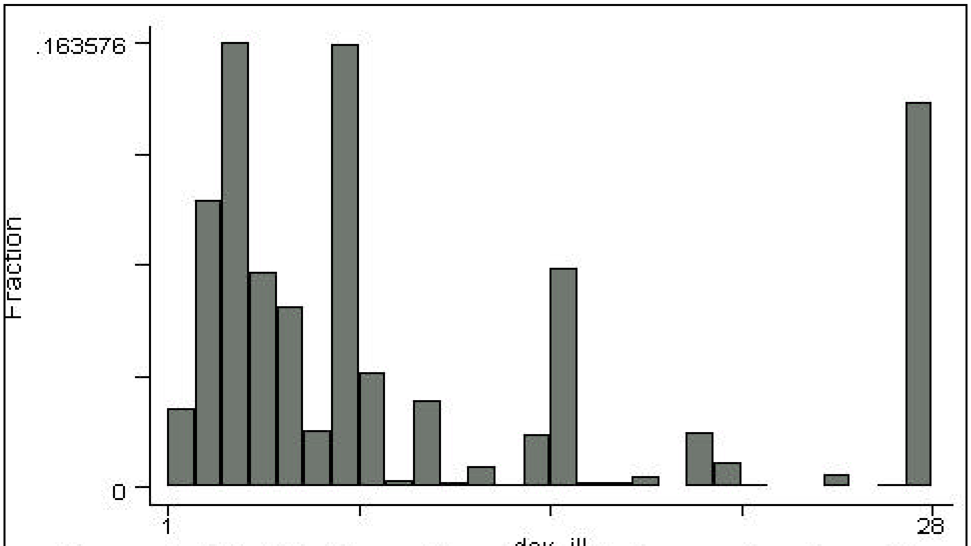
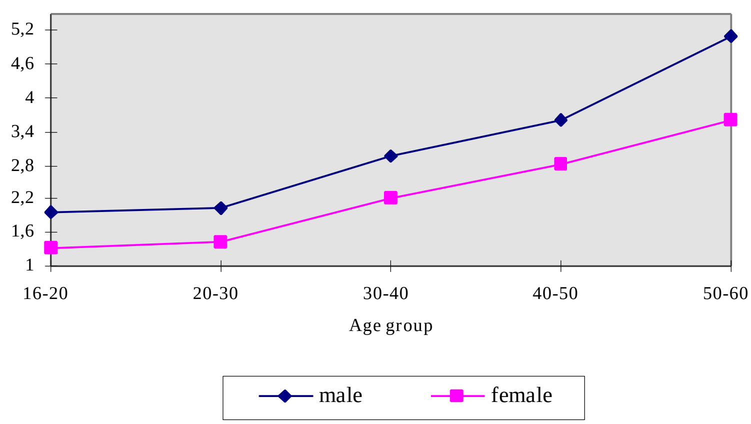


Figure 2. Distribution of the truncated sample - days-ill

As the main question in this paper is the connection between health and productivity, we must discuss the potential impact of particularities in the Peruvian labor market. The Peruvian economy has a considerable fraction of its labor force in self-employment, and this fraction has been growing through time. Table 1 in Appendix B shows descriptive statistics for the full sample of males and females and also for those individuals involved in each sector.

In the 1994 PLSMS, about 33% of the individuals aged 16 to 60 were involved in the wage sector, while almost 27% was performing self-employed activities. These proportions represent 56% and 44% of all working individuals, respectively. The rest were either unemployed, working as unpaid family workers, or not participating. The proportion of self-employed among those working also differs across genders. Among females the proportion of self-employed workers is 49%, but it is only 41% among males.

Figure 3: Average reported days-ill by age group and gender
(Incluye day_ill=0)



The measure of earnings used in this study is the average hourly wage or net income over the twelve months prior to the time of the interview. This indicator varies significantly between sectors, with the earnings of the self-employed being larger than those of wage earners. When the differences are examined by gender, it appears that self-employed males are better paid than those in the wage sector. Among females, however, wage earners are better paid, but the difference is not significant. This may reflect a gender gap in the accumulation of human capital, since self-employed females have four years less education on average. Differences can also be observed in morbidity indicators. The self-employed report the most days ill in the last four weeks, followed by those who are not participating in the labor force; with wage earners reporting the fewest.

Given the observable differences among individuals in different employment sectors, one could easily suspect of other unobservable differences among these individuals that could generate some types of self-selection. Therefore, we need to control for such selectivity in order to estimate wage equations for each sector. Considering these issues, the next section discusses the econometric techniques used to address some of the potential biases in this study.

4. The Econometric Method and Strategy

In this section, we discuss the econometric methods used to estimate the entire system of regressions, including the (negative) health status equation, the participation equation, and the wage equation. First, we explain the model specification for each of these equations. Then, we discuss the simultaneity issues that complicate the estimation, and finally present the strategy to estimate the desired unbiased parameters.

The Health Status Equation

First, we specify a latent health variable h^* that is determined by a linear specification of equation (4) from section 2. That is, $h_i^* = H_i^{'} b + m_i$, where $H_i = \left(P_c, P_y, S, F, Z^i, u^i\right)$. As we observe only the proxy, number of days ill (h), the relationship between h^* and h is given by the following conditions:

$$
\begin{aligned}
h_i &= 0 \quad && if\ h_i^* = H_i^{'} b + m_i \leq 0 \\
& h_i^* && if\ c > h_i^* = H_i^{'} b + m_i > 0 \qquad (6) \\
& 28 && if\, h_i^* = H_i^{'} b + m_i \geq c
\end{aligned}
$$

The double censoring of the proxy h requires that we estimate h^* through a two-limit tobit model. Following Lee (1981), we then obtain the predicted values $\hat{h}_i = H_i^{'} \hat{b}$ and insert this constructed variable in the wage equation[5]. To identify health in the wage equation, we use food prices, housing infrastructure and indirect health cost variables in H, according to equation (4) in section 2. They are both assumed to be highly correlated with the individuals' health status and uncorrelated with individual wages. Both assumptions will be tested with appropriate specification tests.

According to (4), self-reported individual health status depends on prices of consumption and health-related goods, household full income, access to sanitary infrastructure and individual observed and

---

[5]. Alternatively, we could have used $max\left\{0, H'\hat{b}\right\}$ in the wage equation, an alternative which is discussed below.

unobserved characteristics. Here, we will use self-reported morbidity events in the 4 weeks previous to the application to the survey as a proxy for (negative) health status.

For measures of full income *S*, we take the household's productive and non-productive assets. As non-productive assets, we include the value of the house and other durables. As measure of productive assets, we include a dummy for the operation of a household business. We are not using the reported profits from household business since there is a considerable amount of reporting error in the data, something already observed in developing countries (Thomas and Strauss, 1997).

Individual observed characteristics include the age of the individual and his/her education, among others. As measures of sanitary infrastructure *F*, we use the existence of in-house connections for potable water and sewerage, but also whether the house has proper floors, walls and ceiling. Dirt floors tend to provide less healthy environments. We also proxied prices of health services with the costs of having access to those services, e.g., the distance to the health center and the waiting time before receiving attention. Since these indicators are reported only by those who were sick and received medical attention, we use district averages in the regression. Given intra-district correlation of errors, we use the corrected covariance matrix suggested by Moulton (1986, 1990).

The Sector Choice Model

As discussed in section 3, the estimated wage equations in each sector must be corrected for potential selectivity biases. In this paper, we model the sectoral choice as a multinomial logit (MNL) model, and then we construct the correction terms for the selection indicator following Lee (1982,1983). Specifically, we identify three employment status choices (0=non-participation, 1=wage-earners and 2=self-employment)[6], and estimate a multinomial logit in which non-participation is the base category, and where the underlying latent variable is a linear function of a set of observable characteristics of the individual, B_i. That is,

$$b_{Si}^* = B_i'\,g_S + s_S n_{Si}, \qquad (7)$$

where v_{Si} is an unobserved error term and σ_S the variance of this term.[7] Then, individual *i* chooses to be in sector *k* if

$$b_{ki}^* = \max\{b_{Si}^*\}, \quad \text{for } S = 0, 1, 2 \qquad (8)$$

This choice model has a utility interpretation where b_{ki}^* is interpreted as the indirect utility for individual *i* of choosing sector *S*. Then, a utility maximizing individual will behave as in equation (8).

---

6. The self-employment sector includes both, independent workers as well as firm owners.

7 Recall that binary and multinomial choice models identify the coefficients up to a scale, usually assuming that the variance is equal to one.

This equation included controls for individual, household and community characteristics. The variables used for identification of participation in the labor force are community labor market characteristics and household assets and other non-labor income, as defined for the health status equation.

The wage equation

Given the employment status choice, each individual has an associated wage in each sector, which is described by the following equation

$$\ln w_{si} = \delta_{S0} + \delta_{S1} E_{ij} + \delta_{S2} A_{ij} + \delta_{S3} \hat{h}_{ij} + \delta_{S4} L_j + \varepsilon_{Si} \qquad (9)$$

for S = 1, 2, since no equation is estimated for the non-participants. The earnings equations for wage-earners and self-employed are estimated taking into account the necessary selection correction, analogous to the Heckman two-step estimation procedure (see Lee, 1983).

The selection correction term variable is identified in the wage equation through the inclusion of household assets and non-labor income in *B*, which are assumed to be highly correlated with individual participation in the wage labor market, but uncorrelated with individual wages. Both assumptions will be tested with appropriate Wald tests.

The controls included potential experience and its square, defined as the number of years after leaving school. The terms for experience and education in (9) could be included in different ways. Age and health, for instance, could include a linear as well as a quadratic term. Schooling is included in a spline specification, in order to capture differential returns to education at each education level. Three terms are included. First, the plane number of years of schooling which is interpreted as the returns to primary school. The second term, secondary spline, is the product of a dummy equals to one if the individual reached some level of secondary education, multiplied by the number of years. It captures the *additional* returns to secondary school, that is between 7 and 11 years of education. The third term, higher spline, is defined analogously for post-secondary schooling and captures the additional returns to higher education that are accumulated to those obtained in primary and secondary education. Consequently, the returns to an additional year at the university, for instance, would be obtained by adding the value of the three coefficients..

Firm size controls were also included in the earnings equations for wage earners, following the motivating work by Anderson (1998). A firm is classified as a micro firm if it has between six and ten workers. Those having between eleven and twenty are named small, those with more than twenty but no more than 200 workers are medium sized firms, and those with more than 200 workers are large firms[8]. The omitted category is the firm which has between one and five workers. Even though employer size-

---

[8]. Our definitions of medium and large firms are chosen as in Anderson (1998) for comparability purposes.

wage effects are not totally explained by unionization or other regulatory compliances, nor higher workers quality, employer size-wage premium is still found important in developing countries and in Peru, in particular (Anderson, 1998)[9]. In order to capture the potential employer size-wage premium we included indicators for the four firm size categories described above: micro, small, medium and large firms. We also included a control for the degree of market arrangements in each province that may affect the hourly payment in either wage or self-employment.

The community variables included in the model are the local unemployment rate and the fraction of hired labor in the locality. For each province, we use the share of labor days worked in exchange of a wage as the indicator of local labor market development. We presume that when local labor markets are more developed, the number of labor days worked for a wage rises relative to the number of labor days worked as self-employed and unpaid family members. Additionally, we include the size of the firm under the presumption that larger firms tend to pay higher wages (see Anderson, 1998). The selection correction term is included under IMR. Regional dummies control for geographic differences among the regions compared to Lima (omitted category).

Some Empirical Issues in the Estimation of the Effects of Health Status on Productivity

In this last sub-section, we expand on the issues that must be addressed to properly measure the effects of health status on productivity. First, a discussion of the measures for health status and productivity examines the advantages of the variables used and the potential problems involved. Second, the endogeneity of health status as a source of bias is discussed to propose an instrumental variable estimator.

In order to deal with the endogeneity issue discussed in section 3, this health status measure is instrumented with local community sanitary and household infrastructure. The strong assumption with these instruments is that the unobserved individual characteristics underlying the determinants of health status are not correlated with location decisions (i.e., migration), once other household and individual characteristics have been controlled for. This is not an unusual assumption, since other papers (Schultz and Tansel, 1997; Thomas and Strauss, 1997) use the same strategy to instrument health status with community prices and infrastructure. The identification of the model is possible due to the inclusion of variables in the health status equation that are not included in the wage equation such as the community infrastructure variables described before.

Another issue in the estimation arises from the fact that in the 1994 PLSMS, only about one third of the sample reported a positive number of days sick during the last four weeks. This relatively small

---

[9]. We did not include firm size in the earnings equation for the self-employed, understanding that we would

fraction of uncensored information on health represents a problem when estimating the relationship between health and productivity because such health indicator does not allow to distinguish the health status differences among those that did not report any sickness. In this paper we circumvent this problem by using the predicted estimate from the tobit model as an index for the latent health variable. By using the uncensored estimated index we are assuming that a *healthier* person has a lower index (more negative) and try to evaluate whether this implies higher productivity for the individual. We use this estimated index in the second stage wage equation.

Before explaining the results, some caveats regarding the econometric strategy should be mentioned. First, unobserved heterogeneity in the health status regression could cause three main problems. If those unobserved characteristics are correlated with the regressors, the estimates might be biased. On the other hand, if these unobserved components are not correlated with observed characteristics, heteroskedasticity patterns might appear. Even though this problem does not affect consistency in linear models (OLS) , it is a potential source of inconsistency in nonlinear models (see Maddala, 1981 p. 179). The conditional expectation, however, is still consistent, and this is the instrument used in the second stage estimation. Second, the tobit equation actually identifies the parameters of the model based on the uncensored section of the data. The conditional expectation used in the second stage is based on those estimates to predict the latent health status variable for the entire working sample. The assumption is that the estimated parameters are common for all individuals. Third, to obtain the conditional expectation of the health status through the Tobit estimates a normality assumption is imposed. While this is a direct approach to this problem, the conditional expectation loses any non-normal behavior that could be present in the latent health variable. If these data peculiarities are important, a more flexible approach should be examined.

It must be mentioned, however, that these caveats could not be circumvented when the censored portion is relatively large, as in our case. For instance, semi-parametric estimators for censored models that are robust to heteroskedasticity or other mis-specifications (Powell 1984, 1986) require a small portion of censored observations. Similarly, a non-parametric approach would require uncensored information to create a conditional expectation of the health variable. In sum, given the limitations of the data, distributional assumptions have to be imposed in order to extend the analysis to a larger part of the sample. In a rather practical approach, Schultz and Tansel (1997) imposed a linear relationship to instrument the number of days disabled using an OLS estimation. This approach, however, does not distinguish the amount of information underlying an unobserved health status (zero disabled days) from

---

need a different interpretation for those estimates, and that no previous empirical work would provide as a reference.

the variation among those with observed morbidity. The tobit approach used here is more informative since it properly weights information from the censored and uncensored parts of the sample.

5. The Econometric Results

In this section we present the results for the health status equation, the sectoral choice model, and the wage equation. The analysis of the results of the wage equation include a discussion of the results for the overall sample, by age groups, across the wage distribution, and by public/private employer.

The Health Status Equations: A Two-limit Tobit Model

As indicated in section 4, we estimate the health status equation as a function of individual characteristics (age, schooling), household characteristics (such as assets and house infrastructure), cost of health services (distance and waiting time for medical attention), as well as some regional and date of interview controls. The individual characteristics include age, education, and rural background through a dummy that indicates who migrated from rural areas. Some interaction terms are also included. First, the interaction between schooling and age and the implicit health prices. Second, the interaction between rural background and age.

Table 1 shows the results of the estimation of the tobit model for males and females[10]. Since days-ill is actually negatively correlated with unobserved health status, a positive coefficient means that the corresponding variable affects negatively the individual's health status. For instance, the positive coefficient for age indicates that the health of individuals deteriorates with age. The negative effect for the housing infrastructure variables indicates that individuals living in houses with appropriate sanitary facilities tend to be healthier. It is important to notice the negative coefficient found for the distance to the closest health facility, which would indicate that the more accessible the health facility, the more aware is the individual of their health status. This result shows the importance of the reporting bias associated to the indicator of health status used here. The same is true for the positive coefficient found for education, an effect that is discussed below.

---

[10]. In table A.1 of appendix A, we report the regressions using disabled days. Table A.2 in appendix A compares the probit with the tobit estimation, and also includes an interval regression using weeks-ill.

Table 1
Health Regression for Peruvian adults by gender and location: A two-limit
(t statistics in parentheses)

Variables	Males		Females	
Individual Characteristics	*83.59*	**	*69.6*	**
1 - Age	0.675	**	0.243	**
	(5.67)		(2.96)	
2 - Years of schooling	0.720		-0.735	
	(1.18)		-(1.51)	
3 - Schooling x Age (x 10^{-2})	-3.996	**	-0.066	
	-(3.51)		-(0.09)	
4 - Rural migrant	11.539	**	-1.982	
	(3.55)		-(0.59)	
5 - Rural migrant x Age	-0.236	**	0.076	
	-(3.04)		(0.88)	
Household Assets	*5.98*		*0.77*	
6 - Non-Labor Income	0.132		0.056	
	(1.24)		(0.64)	
7 - Nonproductive assets	-0.093		0.024	
	-(0.51)		(0.14)	
8 - Home business	2.471	**	0.572	
	(2.13)		(0.59)	
Housing Infrastructure	*6.12*	**	*7.19*	**
9 - Adequate ceiling	-2.707	**	-2.864	**
	-(2.29)		-(2.63)	
10 Adequate floors	-0.533		0.394	
	-(0.45)		(0.33)	
Health Infrastructure	*7.84*	*	*4.45*	
11 Distance time to health service	-13.375	**	-9.651	*
	-(1.80)		-(1.64)	
12 Wait. time to medical attention	-0.015		-2.649	
	-(0.01)		-(1.25)	
13 Dist. time to health service x schooling	1.314	**	0.681	
	(2.40)		(1.33)	
14 Wait. time to medical attention x schooling	0.123		0.246	
	(0.50)		(1.20)	
Food Prices	*5.59*	*	*14.27*	**
15 Potato price	7.265	**	5.087	
	(2.36)		(1.35)	
16 Milk price	3.688		13.330	
	(0.83)		(3.58)	
Log Likelihood	-3399.62		-4863.72	
Global Chi-squared	281.00		387.27	
Number of observations	3083		3486	

(*) Statistically significant at 10% level of confidence.
(**) Statistically significant at 5% level of confidence.
/ A constant and control variables for regions and interview months are included in all regressions, but not reported.

Household permanent income, as proxied by the per capita value of family assets (variables 6 to 8 in table 1), is included in the regression analysis, but none of them appear to be significant. Although, any theoretical framework would imply the need to include some measures of household income, previous studies have not succeeded in obtaining a strong empirical relationship (see, for instance, Schultz and Tansel, 1997). We also tried some other alternative measures such as per capita household expenditures (exogenous) and per capita full income (instrumented) but neither was significant.

The analysis presents consistent and robust results for males, but not so much for females. Nevertheless, the proposed instruments for endogenizing health in the wage equation are strongly significant for both sub-samples. Recall from section 2 that key instrumental variables are the indicators for food prices, housing infrastructure and implicit prices for health inputs. Access to a house with adequate floors and ceilings does increase the health status of adult individuals.

Differences of human capital effects over the life-cycle

It is very interesting to see the interaction between the different forms of human capital, as measured by age, completed years of schooling, and migration experience, over the life cycle (across cohorts). First, even after controlling for other factors, the health of Peruvian adults clearly deteriorates with age, an effect that is stronger for male adults. When included, the quadratic term for age was not found significant, a result consistent with the strong linear correlation observed in figure 3.

Somewhat surprisingly, table 1 seems to indicate that education matters only for males. Nevertheless, when calculating net schooling effects (table 2), we see that this variable is also significant for adult females at particular ages[11]. But there are, indeed, important gender differences in the role of human capital endowments for the determination of individual health status. First, more educated males have more days-ill reported over the four weeks prior to the survey, when they are younger than 33, but the effect turns increasingly negative afterwards. This result differs from those of Strauss et.al., (1993) with anthropometric measures, and would indicate differences in the magnitude of the self-reporting bias over the life cycle. That is, young non-educated males may be less inclined to report illnesses than their less educated counterparts. At older ages, though, the seriousness of the illnesses may eliminate subjectivity in reporting. On the other hand, the schooling effect for females is constant around -0.25 over the life cycle. This implies that the schooling effect for individuals around 30 is larger (more negative) for females than for their male counterparts, but

[11]. Net effects reported in table 2 refer to the variation of the schooling and migration effects due to the interaction terms with age and implicit health prices included in the tobit model of table 1. Table 2 also includes the estimated net effects associated to a regression without the interaction terms for education and implicit health prices. It is interesting to notice that, although, the individual coefficients for education change significantly, that is not so for the estimated net education effects.

much smaller afterwards. These results are corroborated when looking at the regressions by age groups, as reported in table A.3 of appendix A.

Table 2: Net effect on health by gender and age groups

Age	Male W/o	Male Base	Female w/o	Female Base
Net Schooling effect on				
16 years	0.70 **	0.69 *	-0.25	-0.24
	(0.01)	(0.01)	(0.29)	(0.30)
30 years	0.13	0.13	-0.25 *	-0.25 *
	(0.37)	(0.38)	(0.09)	(0.10)
45 years	-0.47 **	-0.47 *	-0.26 *	-0.26 *
	(0.00)	(0.00)	(0.02)	(0.03)
60 years	-1.08 **	-1.07 *	-0.27	-0.27
	(0.00)	(0.00)	(0.12)	(0.12)
Net Migrant effect on Health				
16 years	7.59 **	7.76 *	-0.63	-0.77
	(0.00)	(0.00)	(0.77)	(0.72)
30 years	4.44 **	4.46 *	0.40	0.30
	(0.00)	(0.00)	(0.74)	(0.80)
45 years	1.06	0.92	1.50	1.44
	(0.40)	(0.47)	(0.18)	(0.20)
60 years	-2.31	-2.62	2.60	2.58
	(0.25)	(0.19)	(0.22)	(0.23)

/ Results in first and third columns refer to regressions reported in table 1. Results in second and fourth columns are based on a regression that includes interaction terms between schooling and housing Infrastructure. The numbers in parentheses are the corresponding p-values.

The other indicator of human capital included for the regressions for urban areas was the dummy variable indicating whether the individual had migrated from a rural area. The positive coefficient in table 1 indicates that, other things equal, a rural background implies poorer health status for adult males, either due to the poorer previous sanitary environment or economic situation. It is also possible, though, that such an indicator could be capturing some ethnic differences since it is limited to migration events from rural areas where there is a strong concentration of individuals with indigenous background. But, if that were the case, the effect would be persistently negative (or positive) over the life cycle. That is not the case, though, since the coefficient of the interaction term is clearly negative[12]. Table 2 shows that the net effect is indeed negative for males around 50 years old and older. These findings support the interpretation that the migration indicator captures previous sanitary and economic conditions that put migrants in disadvantage with respect to urban

[12]. When regressing by age groups (table A.3 in appendix A), we also find that the migration indicator implies poorer health for young (below 30), but has the opposite effect for older individuals (above 45). For the intermediate age group the effect is still negative, but significantly smaller.

natives when they first move. The other implication is that there is evidence that significant *catching up* in health conditions occurs as time passes[13].

Employment Status Choice: Non-participation, Wage and Self-employment

The potential selection bias in the estimation of the earning equations in both sectors is corrected here using the MNL-OLS two-stage correction method proposed by Lee (1983). In Table 3, MNL sectoral choice estimates for males and females are shown. Education is controlled in the sectoral choice model by including three dummies for each level of formal education, and interactions of these dummies with the number of years of schooling. This flexible functional form is intended to capture any non-linearities in the effects of education on sectoral labor choice. The results indicate that education increases the participation of individuals in a non linear way. Moreover, as education increases, the likelihood of joining the wage sector increases faster than the self-employment sector. The estimated coefficients for age and age squared indicate an inverted U-shape for males and females, where males peak around 43 years and females around 38 years.

Wealth and income control variables included a measure of assets, non labor income and the existence of a household business. The accumulation of non-productive assets has an important negative effect on joining any sector, but is even stronger for wage employment, probably reflecting some complementarity between household non-productive assets (a stove, for example) and self-employment activities (cooking). The home business variable can be also interpreted as the existence of another income source, which clearly increases the propensity to join the self-employment sector as a firm owner. The results indicate a suggestive difference for males and females. Among males, we do not find such effect, probably because males are most commonly attached to the wage sector. Among females, the effect is negative, consistent with a traditional behavior where males are associated with the wage sector and females attached to home-related activities. Other explanations as a segmented labor market for males and females are not ruled out based on this evidence.

The effect of local unemployment rate is positive for both genders, but weaker for males. Individuals living in provinces with larger unemployment rates tend to participate more, and even more in the wage sector and among females. In other words, provinces with weak economic growth have higher participation of females, possibly as an additional source of household income. The proportion of hired

---

[13]. These results would be consistent with the empirical literature on the convergence of income of migrants and natives in several countries as time passes (see, for instance, Borjas, 1994). We do not analyze this result further, but it can be said that the evidence here does not reject a hypothesis that the *catching up* in income may be at least partly related to the *catching up* in health reported here.

labor in each province does not have a significant effect on the sectoral choice. There is only an imprecise negative effect on the probability of joining the wage sector among females.

Table 3
Participation Equation for Urban Adults: A Multinomial Logit Model
(t-statistics in parentheses)

		Males		Females	
	Variables	Wage-earner	Self-employ.	Wage-earner	Self Employ.
	Individual Characteristics	*493.22* **	*581.56* **	*323.59* **	*279.81* **
1 -	Age	0.594 **	0.702 **	0.228 **	0.393 **
		(18.29)	(18.85)	(8.48)	(14.26)
2 -	Age squared (x 10^{-2})	-0.734 **	-0.824 **	-0.341 **	-0.469 **
		-(16.69)	-(16.65)	-(8.84)	-(12.88)
3 -	Dummy primary school	2.130 **	1.649 **	-0.403	-0.145
		(3.06)	(2.29)	-(0.92)	-(0.45)
4 -	Years of school (Primary)	-0.130	-0.015	-0.041	0.043
		-(1.16)	-(0.13)	-(0.53)	(0.78)
5 -	Dummy secondary school	0.976	1.038	-1.580 **	0.033
		(1.26)	(1.21)	-(2.29)	(0.05)
6 -	Years of school (Secondary)	0.061	0.048	0.142 **	-0.003
		(0.94)	(0.66)	(2.28)	-(0.05)
7 -	Dummy high education	-1.695 *	-0.395	-5.427 **	-1.219 *
		-(1.63)	-(0.33)	-(6.29)	-(1.04)
8 -	Years of school (High)	0.204 **	0.093	0.441 **	0.077 *
		(3.04)	(1.20)	(7.77)	(0.97)
	Household Assets	*28.89* **	*344.27* **	*12.41* **	*611.09* **
9 -	Non-Labor Income (x 10^{-2})	-0.467	-2.770 **	0.688	0.246
		-(0.42)	-(2.12)	(0.77)	(0.25)
10 -	Nonproductive assets	-0.155 **	-0.144 **	-0.046 **	-0.077 **
		-(5.25)	-(4.45)	-(2.27)	-(3.80)
11 -	Home business	-0.007	21.330 **	-0.237 **	20.785 **
		-(0.06)	(18.35)	-(2.44)	(24.67)
	Labor Market Variables	*6.24* **	*1.89*	*7.41* **	*4.96* *
12 -	Unemployment rate (by province)	1.556 *	1.426	2.073 **	1.764 **
		(1.74)	(1.37)	(2.65)	(2.21)
13 -	Hired rate (by province)	0.641	-1.019	-1.384 *	-0.964
		(0.69)	-(0.85)	-(1.85)	-(1.02)
Number of observations		3102		3508	
Log Likelihood		-2290.5		-2770.3	
Global Chi-squared		546.4 **	205851 **	349.6 **	338244 **
Chi-squared of Instruments		33.2 **	798.1 **	17.8 **	1657.9 **

/ See note to table 1.

Using the estimates from the sectoral selection model, we constructed the selection correction terms described above and included it in the wage equation. In the next part, the endogeneity of health status is examined.

Instrumenting health status in the wage equation

Here, we discuss the results of inserting health in the wage equation, using the information obtained from the health status and sectoral choice equations. We find a robust positive effect of health status on productivity and wages, and also some evidence of interaction between health and other human capital variable, education.

In Table 4.a and 4.b, we present different specifications for the log hourly net earnings equations for wage earners and self-employed, respectively. In both cases, we include estimated earnings equations without health (columns 1 and 4), which are discussed first in order to establish the consistency of our specification with previous estimations. For instance, an important finding in terms of the estimated returns to schooling is that they are higher for females and increasing in the level of educational attainment, especially among wage earners. The use of a spline specification for the returns to schooling allows us to find this pattern. As explained in section 4, the first coefficient in column 1 of table 4.a means that each year of primary schooling imply an increase of 5,2% in wages for males. The second coefficient is not significantly different from zero, which means that such return is about the same for each year of secondary school. Finally, the third coefficient means that each year of post-secondary school would increase wages for males in about 13,0% (5,2+1,2+6,6).

Among females, returns to primary and secondary schooling go up to 8,7%, which is 3.5 percentage points above the returns for males. Nevertheless, this gender difference almost vanishes when individuals reach higher education. The results for the self-employed show a very similar pattern although less precise. The return to primary schooling among males is 5,0% and 7,7% for females, reproducing the gender difference observed among wage earners. Returns to higher education increase for self-employed males. Nevertheless, no evidence of such increasing pattern is identified for females.

The increasing pattern found for returns to education is consistent with previous findings by Saavedra (1997) for the Peruvian case[14]. These increasing returns would support the notion that each pool of workers (skilled, unskilled) faces a qualitative different labor market, where there is abundance of people with primary and secondary education and a scarce number of workers with higher education, as it is the case in Peru (see table 1 in appendix B).

In terms of the firm size effect, our results indicate a significant positive effect, consistent with Anderson (1998). Among wage employed males the premium is about 16% for micro firms (6-10 workers), goes up to 23% for small firms (11-20) and to 36% for medium firms (21-200), and remains almost constant afterwards. Among females, the estimated firm size premium shows a similar pattern, although no significant effect is found for micro firms. We did not include firm size controls in the earnings equation for the self-employed, understanding that the transmission mechanism would need to be different.

Tables 4.a and 4.b also show the wage regressions including health, both, as exogenous and instrumented (IV). The inclusion of health as exogenous show no significant effect, and does not change any of the estimates for the control variables, for any gender or sector. This may have different explanations. Endogeneity in the health measure would imply that those with better unobserved skills, who are probably earning higher wages, will report more days sick generating a positive bias in the estimation. On the other hand, a purely random measurement error biases the estimates towards zero. Finally, the censoring that affects our morbidity indicator also reduces substantially the variation in this explanatory variable making it hard to find any precision in the effects. A combination of these factors could lead to the imprecise effects of exogenous health on wages.

When instrumenting health in the wage equations, we do found significant effects of Days Ill on wages in both sectors, although albeit small. Among males, an increase of one unit in our morbidity indicator implies a 1.2% decrease in wages and a 3.1% decrease in hourly net income for the self-employed. One possible explanation for this result is related to sectoral differences in the observability of individual productivity. In the wage sector, the employer cannot observe individual productivity (effort) by looking at the final output, when the production function is stochastic and the external shock are not observed. Even in a deterministic environment, employers face an observability problem if tasks are performed in large teams[15]. The results among females does not show any significant sectoral differences. The effect of our morbidity indicator is –2.4% among wage earners and –2.3% among the self-employed sector.

---

[14]. Saavedra (1997) reports this same pattern for 3 samples of the available PLSMS. However, this result differs from the decreasing pattern reported by Birdsall (1996) for other Latin American countries.

[15]. Actually, we cannot discard the possibility that part of the differences found here would be related to the different nature of contracts in these two sectors, as a result of which, wage-earners receive a lower premium for health. The explanation based on different contractual agreements is discussed further when examining the differences between public and private workers.

Table 4.a

Equation for the Log Hourly Earnings for Urban Wage Earners by Gender: Instrumenting Health

(t-statistics in parentheses)

		Male			Female	
		No health	Exog. Health	IV Health	No health	Exog. Health
	Individual Human Capital Variables	*94.45* **	*70.99* **	*72.43* **	*28.39* **	*22.24* **
1 -	Years of schooling (x 10^{-2})	5.161 **	5.186 **	4.757 **	8.709 **	8.753 **
		(2.31)	(2.33)	(2.13)	(2.94)	(2.94)
2 -	Years of schooling minus 6 (x 10^{-2})	1.206	1.177	1.753	-2.295	-2.657
		(0.42)	(0.41)	(0.60)	-(0.53)	-(0.61)
3 -	Years of schooling minus 12 (x 10^{-2})	6.609 **	6.607 **	6.384 **	7.020 **	7.298 **
		(3.48)	(3.48)	(3.37)	(2.35)	(2.44)
4 -	# of days sick-instrumented (x 10^{-2})		0.081	-1.216 **		-0.552
			(0.20)	-(2.23)		-(1.28)
	Other Individual Characteristics	*37.24* **	*36.69* **	*33.72* **	*22.91* **	*23.03* **
5 -	Potential Experience	0.029 **	0.029 **	0.029 **	0.030 **	0.030 **
		(4.86)	(4.87)	(4.97)	(3.79)	(3.80)
6 -	Potential Experience squared (x 10^{-3})	-0.313 **	-0.314 **	-0.268 **	-0.279	-0.277
		-(2.44)	-(2.45)	-(2.08)	-(1.36)	-(1.35)
	Firm Size and Local Labor Market	*12.20* **	*12.19* **	*11.91* **	*7.81*	*7.81* **
7 -	Micro scale firm	0.156 **	0.156 **	0.155 **	5.4E-03	7.6E-03
		(2.76)	(2.77)	(2.74)	(0.07)	(0.10)
8 -	Small scale firm	0.232 **	0.231 **	0.233 **	0.213 **	0.216 **
		(3.69)	(3.67)	(3.71)	(2.62)	(2.66)
9 -	Medium scale firm	0.355 **	0.355 **	0.347 **	0.327 **	0.328 **
		(7.20)	(7.21)	(7.04)	(4.62)	(4.63)
10 -	Large scale firm	0.399 **	0.400 **	0.401 **	0.371 **	0.375 **
		(5.36)	(5.36)	(5.39)	(3.79)	(3.83)

(continuation of table 4.a)

	Male			Female	
	No health	Exog. Health	IV Health	No health	Exog. Health
11 - Hired rate (by cluster)	-0.019	-0.018	-0.052	0.294 *	0.289 *
	-(0.16)	-(0.15)	-(0.43)	(1.76)	(1.74)
12 - Selection term	-0.042	-0.042	-0.020	-0.072	-0.075
	-(0.81)	-(0.80)	-(0.38)	-(0.56)	-(0.59)
Number of observations	1543	1543	1543	844	844
F-test	39.45 **	37.17 **	37.52 **	33.00	32.22 **
R-squared	0.281	0.281	0.284	0.351	0.352
Exogeneity Test (Hausman)			11.79		
			(0.00)		
Over-identification Test			19.37 **		
			(0.02)		

(*) Statistically significant at 10% level of confidence.

(**) Statistically significant at 5% level of confidence.

/ A constant and regional control variables are included in all regressions, but not reported. Numbers in italics to the right of the labels of variables refer to the Wald statistics for the joint significance tests.

Table 4.b

Equation for the Log Hourly Earnings for Urban Self Employed by Gender: Instrumenting Health

(t-student in parentheses)

		Male			Female	
		No health	Exo Health	IV Health	No health	Exo Health
	Individual Human Capital Variables	*54.04* **	*40.85* **	*42.77* **	*18.66* **	*14.01* **
1 -	Years of schooling (x 10^{-2})	5.006	4.901 *	3.527	7.665 **	7.576 **
		(1.52)	(1.50)	(1.06)	(3.35)	(3.30)
2 -	Years of schooling minus 6 (x 10^{-2})	4.819	4.901	5.829	-0.464	-0.309
		(1.14)	(1.17)	(1.39)	-(0.13)	-(0.09)
3 -	Years of schooling minus 12 (x 10^{-2})	5.553 **	5.520 **	5.647 *	1.420	1.336
		(1.83)	(1.82)	(1.87)	(0.36)	(0.34)
4 -	# of days sick-instrumented (x 10^{-2})		-0.761	-3.147 **		-0.366
			-(1.23)	-(3.24)		-(0.71)
	Other Individual Characteristics	*1.14*	*1.29*	*3.85* **	*3.82* **	*3.97* **
5 -	Potential Experience (x 10^{-3})	3.698	2.780	-12.657	7.010	6.699
		(0.19)	(0.14)	-(0.61)	(0.28)	(0.27)
6 -	Potential Experience squared (x 10^{-3})	0.067	0.090	0.440	0.063	0.072
		(0.22)	(0.30)	(1.30)	(0.16)	(0.18)
	Local Labor Market Characteristics	*0.57*	*0.72*	*0.27*	*3.05* **	*2.92* **
7 -	Hired rate (by cluster)	0.157	0.176	0.108	0.425 *	0.414 *
		(0.75)	(0.85)	(0.52)	(1.75)	(1.71)
8 -	Selection term	-0.418	-0.436	-0.675 **	-0.231	-0.238
		-(1.49)	-(1.55)	-(2.20)	-(0.78)	-(0.81)
Number of observations		1144	1144	1144	843	843
F-test		19.54 **	18.43 **	18.40 **	9.55 **	8.82 **
R-squared		0.162	0.164	0.170	0.123	0.124
Exogeneity Test (Hausman)				10.08 **		
				(0.00)		
Over-identification Test				18.34 **		
				(0.03)		

/ See notes to table 4.a

The inclusion of the instrumented morbidity indicator also has also some implications for the estimated returns to education. In particular, estimated returns to primary education are sligthly smaller , although, no significant difference seem to be found for individuals with higher education. Returns to primary education drop from 5.2% to 4.8% in the wage sector. The drop is even stronger among the self-employed (from 4.9% to 3.5%). The effect of including health for those with higher education is significantly smaller in both sectors. For instance, in the wage sector, returns to schooling drop only from 11.6% to 11.2%.

The returns to schooling among females, though, are not affected by the inclusion of health status as it did to those corresponding to males. This could suggest that the interaction between the two dimensions of human capital, education and health, is more important among males, and this could be associated to types of work in which health is as important as education. For example, in jobs where physical tasks are demanded, both education and good health would affect productivity, while in other jobs, education may play a more important role. It could be argued that this result would be related to some sectoral sorting of individuals due to some unobserved characteristics. Nevertheless, the estimated selection correction coefficients suggest an insignificant selection problem between these sectors, with and without health in the wage equation (see the coefficients for the inverse Mills ratio (IMR) term in tables 4.a and 4.b).

Although the discussed IV results seem to be consistent with the stories of endogeneity and measurement error associated to our morbidity indicator, it is still important to examine some diagnostic statistics. A Durbin-Hausman-Wu specification test and an over-identification test are shown at the bottom of tables 4.a and 4.b. The Hausman specification tests suggest that the endogeneity of health status is important in the sense that the coefficient for the health indicator from the OLS estimates is significantly different from that of the IV estimates. This can be clearly observed among all males, but only for those females in the wage sector.

On the other hand, to test the validity of our instruments we used a regression-based over-identification test suggested by Hausman (1983), and not the Basmann version. The results indicate that the over-identification restriction is rejected at the 5% level for males in both sectors and for female wage-earners. These values, however, are close to the critical ones for both genders. As others have mentioned (Thomas and Strauss, 1997) over-identification and exogeneity test results must be taken carefully when the identifying instruments are weak. In this case, the corresponding pseudo-R2 from the Tobit estimates for the health equation are about 0.039 for males and 0.029 for females. The low correlation between instruments and health and the censoring in the health variable make the interpretation of these tests difficult. Moreover, according to Staiger and Stock (1996) any version of the over-identification tests, regression-based or Basmann, would be identically affected by the weak instruments, so the choice between them made no difference.

Although the connection between health status and productivity-wages is shown to be significant in this section, it is hard to be conclusive about any particular explanation for the observed sectoral differences. In the next sections, we examine these effects across different population groups. For example, if older populations are more likely to be sick, the wages of those older individuals may have a different connection with their health status. Moreover, if the interaction between age, education and health is important, as previously shown, then this interaction might be especially important when related to hourly earnings. Given the weak overall regressions for females, we emphasize the discussion of the results for males. Also, since the selection terms are not found to be significant, the rest of the analysis omits that variable and focus only on the endogeneity of health.

Evidence by Age Groups

In the section that analyzed health determinants it was found that health outcomes had a different pattern across different ages. Here, the effects of health on wages are examined by age groups, looking whether more vulnerable groups of the population have a stronger link between health and wages or whether age-groups associated with more physical work had a different effect. Table 5 presents the wage equation estimates for wage earners and self-employed. For each sector, the sample was divided in two age groups: those aged 16 to 25 years and those aged 26 and older.

We find that, although the size of the health effect may not be robust across age groups, the previously reported finding of larger effects among the self-employed persists. We find a precise negative effect among those workers 26 or older of about -2%, while no significant effect was found among those younger workers. A very similar pattern is found among the self-employed, but the effects on the wages of older individuals is even stronger: -4.3%. Among females, the pattern is loosely the same, although, the results are precise only for older wage earners. Clearly, then, it is not the younger group the one behind the precise and negative effect found on the mean. There are several explanations for this result. First, younger workers are less likely to be sick, hence, health effects on wages are harder to capture and this may be a reason for the large standard errors in the estimates. Second, and more importantly, the reporting bias identified to be stronger among the young, increases the imprecision on the estimates for this group. Probably due to a combination of these factors, the stronger effects are found among the older group. Nevertheless, the radical changes in the estimated coefficients for the younger would actually suggest the inability of the model to capture the variability in their wages and productivity.

Table 5
Equation for the Log Hourly Earnings for Urban Workers by Age Groups
(t-student in parentheses)

		Wage Earners		Self Employed	
		< 26 years	>= 26 years	< 26 years	>= 26 years
	Males				
1 -	Years of schooling (x 10^{-2})	6.161	4.029 *	7.201	4.964
		(0.76)	(1.72)	(0.82)	(1.49)
2 -	Years of schooling minus 6 (x 10^{-2})	-3.635	2.722	7.524	3.081
		-(0.38)	(0.87)	(0.62)	(0.70)
3 -	Years of schooling minus 12 (x 10^{-2})	7.255 *	5.854 **	0.140	5.321 *
		(1.78)	(2.53)	(0.02)	(1.64)
4 -	# of days sick-instrumented (x 10^{-2})	1.203	-2.008 **	0.124	-4.274 **
		(1.12)	-(2.84)	(0.06)	-(3.95)
5 -	Potential Experience (x 10^{-3})	-0.451	25.521 **	-13.668	4.804
		-(0.01)	(2.78)	-(0.20)	(0.32)
6 -	Potential Experience squared (x 10^{-3})	0.600	-0.175	3.471	0.253
		(0.26)	-(0.97)	(0.80)	(0.88)
Number of observations		443	1100	191	953
F-test		4.04 **	27.16 **	6.24 **	14.42 **
R-squared		0.115	0.281	0.171	0.161
	Females				
1 -	Years of schooling (x 10^{-2})	22.475 **	5.617 *	-7.501	8.036 **
		(3.70)	(1.76)	-(0.20)	(3.43)
2 -	Years of schooling minus 6 (x 10^{-2})	-19.524 **	3.523	14.596	-2.068
		-(2.29)	(0.81)	(0.34)	-(0.54)
3 -	Years of schooling minus 12 (x 10^{-2})	13.298 **	3.400	0.969	0.784
		(2.35)	(1.01)	(0.08)	(0.19)
4 -	# of days sick-instrumented(x 10^{-2})	-2.510	-2.778 **	-3.484	-2.061
		-(1.22)	-(2.32)	-(0.81)	-(1.44)
5 -	Potential Experience (x 10^{-3})	-43.382	36.388 **	32.729	15.667
		-(0.89)	(2.59)	(0.30)	(0.99)
6 -	Potential Experience squared (x 10^{-3})	6.224 *	-0.287	-2.491	-0.015
		(1.81)	-(0.97)	-(0.31)	-(0.06)
Number of observations		297	547	115	728
F-test		8.90 **	19.84 **	1.97 **	7.93 **
R-squared		0.240	0.347	0.187	0.115

/ see notes to table 4.a.

The age of the individual may not be the only determinant of the connection between health and productivity. If the relation between health and productivity is related to the nature of the tasks performed, an exam by the type of job or the contractual nature of the job would be reasonable. In the next part, the effects of health on wages are distinguished by employment status using the public/private distinction.

Evidence between Public and Private workers

In the previous sub-samples we posed the hypothesis that a stronger health effect was found among self-employed workers since in this sector health (therefore, productivity) and net income are more closely associated. In the wage sector, though, the connection between health and wages depends on the employers' ability to observe productivity. In this section, we examine a rather different hypothesis. Health premiums in the public sector might be lower, or even negligible, than in the private sector because of a different remuneration policy. Public sector employers may not be incentived to monitor their employees' effort due to an objective different from maximizing output, such as keeping employment level constant[16].

Table 6
Equation for the Log Hourly Earnings for Urban Male by type of employment and employer
(t-student in parentheses)

		Males		Females	
		W. vs B.	Pub vs Priv	W. vs B.	Pub vs Priv
	Individual Human Capital Variables	*42.80* **	*75.07* **	*31.34* **	*38.71* **
1 -	Years of schooling (x 10^{-2})	0.048 **	4.513 **	8.815 **	8.140 **
		(2.14)	(2.02)	(3.25)	(2.87)
2 -	Years of schooling minus 6 (x 10^{-2})	0.010	2.154	-4.909	-2.688
		(0.33)	(0.74)	-(1.22)	-(0.67)
3 -	Years of schooling minus 12 (x 10^{-2})	0.057 **	6.695 **	8.670 **	7.815 **
		(2.97)	(3.55)	(2.91)	(2.67)
4 -	# of days sick-instrumented (10^{-2})	-0.015 **	-1.812 **	-2.499 **	-2.813 **
		-(2.52)	-(3.25)	-(2.45)	-(2.57)
	Other Individual Characteristics	*33.04* **	*41.84* **	*26.20* **	*22.56* **
5 -	Potential Experience	0.028 **	0.032 **	0.037 **	0.033 **
		(5.11)	(5.74)	(4.63)	(4.13)
6 -	Potential Experience squared (x 10^{-3})	-0.269 **	-0.304 **	-0.337 *	-0.237
		-(2.15)	-(2.46)	-(1.73)	-(1.22)
	Occupation Variables	*6.22* **	*8.87* **	*8.68* **	*4.63* **
12 -	Number of days sick for blue collar (x 10^{-2})	0.606		1.421	
		(1.08)		(1.38)	
13 -	Blue collar dummy	-0.051		-0.093	
		-(0.47)		-(0.58)	
14 -	Number of days sick of a public worker		0.021 **		0.016 **
			(3.56)		(2.04)
15 -	Public worker dummy		0.220 **		0.336 **
			(2.11)		(2.92)
Number of obs		1543	1543	844	844
F(21, 1521)		37.34	36.82	30.55	32.26
R-squared		0.290	0.291	0.371	0.360
Net effect on health (x 10^{-2})		-0.911	0.305	-1.078	-1.212
		-(1.48)	(0.46)	-(0.81)	-(1.13)

/ see notes to table 4.a.

We examine these differences between the public and private sector. In addition, we distinguish between blue collar and white collar workers, in order to examine the alternative (but not exclusive) argument

16. Moreover, recent labor market deregulation might have strengthen the connection between productivity and wages in Peru, but the added flexibility has been more difficult to implement in the public sector.

that health has a stronger impact among those performing more physically demanding tasks. The estimated equation is[17]:

$$\ln w_{si} = X_{ij} b_S + d_{S3} \hat{h}_{ij} + d_{S5} \hat{h}_{ij} \cdot JOBTYPE_{ij} + d_{S6} JOBTYPE_{ij} + V_{Si} \qquad (11)$$

where $JOBTYPE$ is an indicator that takes the value of 1 if the worker is in the public sector (blue-collar worker). The results for this exercise for wage earners are shown in Table 6. The column labeled 1 shows that no substantial difference exists between the effects among blue and white collared workers, males or females. This evidence is against the physically-demanding-job type of argument, if the blue and white collared workers are reporting without error. The second column shows some significant differences between public and private workers. The estimated effect of (negative) health among private male workers is –1.8% lower productivity, while the net effect among public workers (shown at the bottom of the table) is not different from zero among males. The same pattern is found among females. It must be emphasized that the "public worker" effect is not due to firm size, since that effect has already been controlled for.

Another way to characterize these effects is by examining them across the wage distribution, which is pursued next.

Health effects over the wage distribution

Instead of analyzing different health effects among specific types of employment or employer, here we look at differences in health effects over the wage distribution. It is not unreasonable to think that low paid jobs imply a stronger connection between health and productivity (therefore, wages), for instance, due to more physically demanding tasks. To avoid splitting the sample by wage groups, we follow Buchinsky (1994) and estimate a quantile regression. The quantile regression model is

$$y_i = x_i' b_q + u_{qi} \qquad (12)$$

where q denotes the log wage quantile being estimated, and $x_i' b_q$ denotes the q^{th} quantile. The parameters for the health status variables are interpreted as the effects on the q^{th} quantile of *ln*(wage), not on the mean *ln*(wage) as in the OLS case. Conditional on $x_i' b_q$ this is equivalent to examine the q^{th} quantile of the residual u_{qi}. In this study we evaluate the effects on the 10th, 25th, 50th, 75th, and 90th percentiles. This instrument is useful for several reasons. First, it provides information about the effects at each part of the wage distribution, so it is possible to describe changes in the wage distribution when a

---

[17]. We are aware of the problems associated to selectivity biases. We could try a fully specified participation equation with separate wage regressions for each sector. This method, though, would require additional identifiying instruments which are generally difficult to obtain.

bad health shock occurs, using the Inter-quartile Range (IQR) as an estimate of the wage variance (inequality). Second, the quantile estimator, when computed at the 50th percentile, provides the least absolute deviation (LAD) estimator, which is more robust than the OLS estimator in the presence of outliers.

The results are shown in Tables 7.a to 7.b for wage earners and self-employed, respectively. In each table, the top panel shows the results for males and the bottom for females. Before discussing the health effects it must be observed that returns to primary schooling are larger among the lower wage tail for wage earners. As before, returns to secondary schooling are not any different from primary schooling. On the other hand, the premium to higher education are larger for those around the third quartile. Note also that the returns to potential experience is lowered among the top quartile. Corresponding estimates for the self-employed are less precise.

Table 7.a shows that the effects of bad health shocks on wages are larger for those males at the bottom of the distribution (-2.6% at the 10th percentile). This result implies more vulnerability of lower wage males to negative shocks in health. That is, if a weather shock implies one more day of illness for all males, the wages of the poor would decrease the most. On the more positive note, though, this result would imply that a health policy that result in a positive health shock among male workers would be effective in reducing wage inequality, as measured as the log variance of wages, even if the government's program has some targeting inefficiencies[18]. Among the self-employed, the results show the same decreasing pattern for the health effects up to the median (from –3.8% at the lowest decile to –0.9% at the median). These effects increase for the upper tail, although they are still lower than those found for the lowest decile.

The results for females, however, show the reverse pattern. The largest health effects are found among the top of the log wage distribution. And this finding is similar in both sectors but accentuated among wage earners.

While the result for males somewhat supports the hypothesis of physical tasks associated to lower paid jobs, the results for females only indicates that a different explanation is needed. For example, if high wage earning females do depend more on their health statuses it would be reasonable to expect larger effects around the top of the distribution. Still, there are other possible explanations. One is associated to the limitations of our (negative) health indicator. As lower paid jobs have more observations with uncensored morbidity, measuring health effects is more accurate with them than in the higher wage

[18]. Clearly, the program's impact on wage inequality would be even larger if targeted exclusively among all the poor. The point we make here is that, such a health program would reduce wage inequality, even if some filtering occurs; actually, even if filtering reaches its maximum.

range[19]. In a similar way, self reported measures might be more systematically biased among top earning females than among those with low wages.

Notice that, as in all previous estimations, health effects are found to be larger among the self-employed than among wage earners, but the difference is not clear here for females.

Table 7.a
Equation for the Log Hourly Earnings for Urban Male Wage Earners over the Wage Distribution
(t-student in parentheses)

		10	25	50	75	90
	Males					
1 -	Years of schooling (x 10^{-2})	7.200 **	5.592 *	8.066 **	3.708	0.712
		(2.20)	(1.58)	(3.23)	(1.00)	(0.12)
2 -	Years of schooling minus 6 (x 10^{-2})	-0.723	0.915	-3.488	1.529	5.975
		-(0.18)	(0.22)	-(0.92)	(0.39)	(0.96)
3 -	Years of schooling minus 12 (x 10^{-2})	5.040 **	6.468 **	7.042 **	8.813 **	6.664 **
		(1.83)	(3.59)	(2.94)	(4.11)	(2.31)
4 -	# of days sick-instrumented (x 10^{-2})	-2.619 **	-1.471 **	-1.469 **	-0.866	-0.436
		-(3.79)	-(1.91)	-(2.31)	-(1.32)	-(0.43)
5 -	Potential Experience	0.047 **	0.035 **	0.032 **	0.029 **	0.031 **
		(5.25)	(5.48)	(7.22)	(4.79)	(3.31)
6 -	Potential Experience squared (x 10^{-3})	-0.577	-0.339 **	-0.326 **	-0.268 **	-0.366 *
		-(2.92)	-(2.14)	-(3.13)	-(2.00)	-(1.69)
Number of observations		1543	1543	1543	1543	1543
R-squared		0.167	0.180	0.174	0.174	0.167
	Females					
1 -	Years of schooling (x 10^{-2})	13.754 **	11.812 **	4.833	3.659	9.432 **
		(2.74)	(3.03)	(1.22)	(0.97)	(2.80)
2 -	Years of schooling minus 6 (x 10^{-2})	-9.809	-4.298	1.058	1.071	-6.089
		-(1.43)	-(0.79)	(0.17)	(0.17)	-(1.19)
3 -	Years of schooling minus 12 (x 10^{-2})	12.522 **	6.939 **	8.147 **	7.487 **	7.665 *
		(2.15)	(1.97)	(2.44)	(1.80)	(1.72)
4 -	# of days sick-instrumented (x 10^{-2})	-0.860	-2.134	-2.877 **	-3.868 **	-3.610 **
		-(0.33)	-(1.44)	-(2.46)	-(2.56)	-(2.20)
5 -	Potential Experience	0.043 **	0.039 **	0.032 **	0.035 **	0.022 *
		(2.93)	(4.85)	(4.06)	(2.78)	(1.53)
6 -	Potential Experience squared (x 10^{-3})	-0.499	-0.252	-0.211	-0.199	0.067
		-(1.40)	-(1.10)	-(1.06)	-(0.72)	(0.23)
Number of observations		844	844	844	844	844
R-squared		0.222	0.250	0.255	0.205	0.164

/ see notes to table 4.a.

19. This result could also be generated if public workers were clustered around the upper tail. We checked the robustness of the pattern over the wage distribution to the inclusion of the indicator variables for the public sector. We found that the effects on public workers is equally insignificant and close to zero across the distribution, while the effects on private workers mimic those reported in this section. In sum, it is the private sector workers who are driving the results across the distribution, corroborating the finding that among public sector workers health status has less of an impact on wages.

Table 7.b
Equation for the Log Hourly Earnings for Urban Self Employed by wage distribution
(t-student in parentheses)

		10	25	50	75	90
	Males					
1 -	Years of schooling (x 10^{-2})	1.376	2.559	4.726	7.289 **	9.531
		(0.23)	(0.72)	(1.28)	(2.04)	(1.29)
2 -	Years of schooling minus 6 (x 10^{-2})	8.936	6.995 *	3.813	0.730	0.105
		(0.96)	(1.76)	(0.84)	(0.18)	(0.01)
3 -	Years of schooling minus 12 (x 10^{-2})	-0.682	3.893	4.871	9.116 **	7.479
		-(0.11)	(1.25)	(1.37)	(4.04)	(1.19)
4 -	# of days sick-instrumented (x 10^{-2})	-3.752 **	-2.147 **	-0.930	-2.225 *	-3.160 **
		-(2.00)	-(2.25)	-(1.08)	-(1.82)	-(2.75)
5 -	Potential Experience	0.027 *	0.029 **	0.032 **	0.025 *	0.022
		(1.80)	(2.15)	(3.59)	(1.76)	(0.80)
6 -	Potential Experience squared (x 10^{-3})	-0.321	-0.256	-0.303	0.026	0.201
		-(0.86)	-(0.87)	-(1.53)	(0.08)	(0.41)
Number of observations		1144	1144	1144	1144	1144
R-squared		0.099	0.089	0.104	0.107	0.104
	Females					
1 -	Years of schooling (x 10^{-2})	3.718	6.604 *	7.101 **	5.939 **	8.962 **
		(0.70)	(1.80)	(2.81)	(3.01)	(2.83)
2 -	Years of schooling minus 6 (x 10^{-2})	3.331	-0.114	-1.203	2.026	0.530
		(0.43)	-(0.02)	-(0.35)	(0.64)	(0.09)
3 -	Years of schooling minus 12 (x 10^{-2})	1.808	-3.233	2.636	5.488	1.066
		(0.24)	-(0.48)	(0.52)	(1.13)	(0.18)
4 -	# of days sick-instrumented (x 10^{-2})	-2.047	-2.275	-3.239 **	-3.117 *	-2.016
		-(0.98)	-(0.89)	-(2.13)	-(1.76)	-(0.78)
5 -	Potential Experience	0.038 *	0.044 **	0.029 **	0.029 **	0.018
		(1.73)	(2.66)	(2.34)	(2.78)	(1.18)
6 -	Potential Experience squared (x 10^{-3})	-0.288	-0.460 **	-0.254	-0.105	0.136
		-(0.99)	-(1.90)	-(1.38)	-(0.47)	(0.55)
Number of observations		843	843	843	843	843
R-squared		0.096	0.064	0.067	0.079	0.098

/ see notes to table 4.a.

6. Summary

This report evaluated the effects of health on wages by using different samples and different statistical methods. The findings on health determinants measured as the number of days sick are reasonably interesting for urban males. First, we do not find evidence of non-linearities in the relationship between age and health, a result that differs from several previous studies. A possible explanation is our decision to limit our analysis to adult individuals (ages between 16 and 60). Schooling appears as a very important determinant of health status, and its importance does increase with age. This result appears very robust for this sample. It does differs from previous estimates such as Strauss, et.al., 1993, but that is most likely explained by the use of a self-reported health measure here, compared to the anthropometric measure used by them.

For younger males, a rural background tends to be correlated with poorer health status, apparently due to poorer environmental and/or economic conditions, but the relationship reverses for older individuals. The positive effect of the rural background on the health of males older than 48 would suggest that the key explanatory factor is the economic condition of the individual as a child. Finally, for all the sub-samples (urban/rural, male/female) the group of proposed instruments seem to be significant determinants of the health status of Peruvian individuals, a result that allows us to look rather confidently to the second stage, the effect of health on wages.

Our key result is that health status is positively and robustly associated with wages, especially for urban males. Table 8 summarizes the main results of this paper. First, health effects are found to be stronger for self-employed than among male wage earners, and that was attributed to the difficulty of employers to observe individual productivity. Second, health effects are weak and negligible among younger individuals, emphasizing the importance of reporting biases among the young in the morbidity indicator used here. For these two effects, imprecision in the estimates obtained for females does not allow to add any information on this issue. Third, the larger effects found among self-employed and the negligible effects found among public workers suggest that the effects on wages are larger in jobs where productivity and wages are closely connected, either by observability or by the design of the remuneration policy. Finally, when examining the effects across the wage distribution, the largest effects are found at the bottom of the wage distribution of both wage earners and self-employed males. A result that cannot be explained is that the reverse pattern is found among females.

Non-linear health effects on productivity were not found in previous versions of this paper, and given the censoring degree of the health indicator, the treatment of non-linearities turned out to be a difficult task by itself. The need of more and better information of health status, such as objective measures of health, is evidenced in the paper, since it would enable to separate much of the noise in the self-reported health status variable. A recent paper by Dow, et. al. (1997) evidenced how self-reported health measures would be biased. Using experimental data from the US and Indonesia, they find that self-reported health measures change significantly as a result of changes in, for instance, health service prices, while objective measures remain unchanged.

The key caveat in the analysis presented here is the overall inability to explain the relationship between health, productivity and wages among females. In that sense, it would be important for future research to concentrate on the way women insert themselves into the labor market, that is in the determinants of their participation in the labor market, in the sectors and occupations that they choose, etc. Previous to that, though, it would also be important to validate the findings here with the use of less subjective health indicators, such as anthropometric measures.

Table 8: Summary of Health Effects on Hourly Earnings

	Male		Female	
	Wage-earners	Self-employed	Wage-earners	Self-employed
Full sample	-1.22% **	-3.15% **	-2.41% **	-2.29% *
	-(2.23)	-(3.24)	-(2.32)	-(1.70)
Young	1.20%	0.12%	-2.51%	-3.48% **
	(1.12)	(0.06)	-(1.22)	-(0.81)
Old	-2.01% **	-4.27% **	-2.78% **	-2.06%
	-(2.84)	-(3.95)	-(2.32)	-(1.44)
Quintile 10	-2.62% **	-3.75% **	-0.86%	-2.05%
	-(3.79)	-(2.00)	-(0.33)	-(0.98)
Quintile 25	-1.47% **	-2.15% **	-2.13%	-2.28%
	-(1.91)	-(2.25)	-(1.44)	-(0.89)
Quintile 50 (median)	-1.47% **	-0.93%	-2.88% **	-3.24% **
	-(2.31)	-(1.08)	-(2.46)	-(2.13)
Quintile 75	-0.87%	-2.23% *	-3.87% **	-3.12% *
	-(1.32)	-(1.82)	-(2.56)	-(1.76)
Quintile 90	-0.44%	-3.16% **	-3.61% **	-2.02%
	-(0.43)	-(2.75)	-(2.20)	-(0.78)
Public	0.30%		-1.21%	
	(0.46)		-(1.13)	
Private	-1.81% **		-2.81% **	
	-(3.25)		-(2.57)	

/ Numbers in parentheses are the t-statistics.

This paper, however, allowed other interesting findings regarding the relationship between health and education. Returns to primary schooling were slightly reduced when instrumented health measures were added into the wage equation, suggesting that, when health is omitted, part of the educational effect on wages is due to the increased health and its associated productivity. This finding offers a window of opportunity for public health policy. The potential effects of better health are significant, but are especially large among those less endowed with human capital and hence earning lower wages. In this sense, public health policies might have effects not considered in traditional program evaluations that should be accounted for to support these programs. Moreover, if policy-makers consider health as a relevant human capital measure, health packages could be considered as part of the programs designed to reduce poverty or improve the living conditions. Overweighing education as the single human capital variable and not taking proper care of health outcomes may lead to wrong targeting of development and income programs for the less favored individuals in Peru.

7. References

Amemiya, T. (1985). *Advanced Econometrics*. Cambridge, Mass., Harvard University Press.

Anderson Schaffner, Julie (1998) "Premiums to Employment in Larger Establishments: Evidence from Peru," Journal of Development Economics, 55 (1), pp. 81-113, February.

Becker, Gary (1965). "A Theory of the Allocation of Time". *The Economic Journal* LXXX (200): 493-517, September.

Behrman, Jere and A. Deolalikar (1988). "Health and Nutrition". In Chenery H. and T. N. Srinivasan (eds.), *Handbook of Development Economics* 1: 631-711. New York, North Holland.

Birdsall, Nancy (1996). "Public Spending on Higher Education in Developing Countries: Too Much or Too Little?". In *Economics of Education Review* 15 (4): 407-419, October.

Borjas, George (1994). *"Assimilation and Changes in Cohort Quality Revisited: What Happened to Inmigrant Earnings in the 1980s"*. NBER Working Paper # 4866, Cambridge, September.

Buchinsky, Moshe (1994). "Changes in the U.S. wage structure 1963-1987: Application of quantile regression". In *Econometrica* 62 (2): 405-458, March.

Dow, W., P. Gertler, R-F Schoeni, J. Strauss and D. Thomas (1997). *"Health Care Prices, Health and Labor Outcomes: Experimental Evidence"*. Labor and Population Program, RAND Corporation.

Grosh, Margaret and Paul Gleewe (1995). *"A Guide to Living Standard Measurement Study Surveys and Their Data Sets"*. The World Bank, Living Standard Measurement Survey Working Paper # 120, Washington, DC., September.

Hausman, J. (1983) "Specification and Estimation of Simultaneous Equation Models." In Z. Griliches, and M. Intriligator, eds., *Handbook of Econometrics*, Amsterdam: North Holland.

Lee, Lung-Fei (1983) "Generalized Econometric Models with Selectivity," *Econometrica*, 51 (3): 507-512, May.

.... (1982). "Some approaches to the Correction of Selectivity Bias," *Review of Economic Studies*, 49:355-372, .

.... (1981). "Simultaneous Equation Models with Discrete and Censored Dependent Variables". In Manski and McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, Boston, MIT Press.

Maddala, G.S. (1983). "Limited-dependent and qualitative variables in econometrics". New York, Cambridge University Press.

Mincer, Jacob (1962). "On-the-job Training: Costs, Returns and Some Implications". In *Journal of Political Economy* 70 (5): s50-s79, October.

Mincer, J. and S. Polachek (1974). "Family Investments in Human Capital: Earnings of Women". In *Journal of Political Economy* 82(2), part2: S76-S108.

Moulton, Brent (1986): "Random Group Effects and the Precision of Regression Estimates", *Journal of Econometrics* 32: 385-397.

-------------- (1990): "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units", *The Review of Economics and Statistics*: .

Pitt, Mark and Mark Rosenzweig (1986). "Agricultural Prices, Food Consumption, and the Health and Productivity of Indonesian Farmers". In Singh, Squire and Strauss (eds), "*Agricultural Household Models: Extensions, Applications and Policy*". A World Bank Publication, John Hopkins University Press, Baltimore, February.

Powell, James (1986). "Symmetrically Trimmed Least Squares Estimation of Tobit Models". In *Econometrica* 54 (6): 1435-1460, November.

-------------- (1984). "Least Absolute Deviations Estimation for the Censored Regression Model". In *Journal of Econometrics* 25(?): 303-325, .

Saavedra, Jaime (1997). "Quiénes ganan y Quiénes Pierden con una Reforma Estructural: Cambios en la Dispersión de Ingresos Según Educación, Experiencia y Género en el Perú Urbano". In *Notas para el Debate* # 14: 9-77, August.

Schultz, Paul T. and Aysit Tansel (1997). "Wage and labor supply effects of illness in Côte d'Ivoire and Ghana: instrumental variable estimates for days disabled". In *Journal of Development Economics*, 53 (2), pp. 251-286, August.

Staiger, Douglas and James Stock (1997). "Instrumental Variables Regression with Weak Instruments". In *Econometrica* 65 (3): 557-586, May.

Strauss, John and Duncan Thomas (1995). "Human resources: Empirical modeling of household and family decisions". In J. Behrman and T.N. Srinivasan (eds.), *Handbook of Development Economics*, Vol. 3: 1883-2023, North-Holland Press, Amsterdam.

Strauss, John; Paul Gertler, Omar Rahman and Kristin Fox (1993). "Gender and Life-Cycle Differentials in the Patterns and Determinants of Adult Health". In *The Journal of Human Resources*28(4): 791-837, Fall.

Thomas, Duncan and John Strauss (1997). "Health and wages: Evidence on men and women in urban Brazil". In *Journal of Econometrics* 77 (1): 159-186, March.

Valdivia, Martín and Miguel Robles (1997). "Decisiones Laborales en las Economías Rurales del Perú". In *Notas para el Debate* # 14: 79-131, GRADE, Lima, September.

Appendix A

Table A.1
Health Determinants for urban adults by gender: Illness vs. Disability
(t-statistics in parentheses)

		Male		Females	
	Variables	illness	Disability	Illness	disability
	Individual Characteristics	*84.44* **	*21.7* **	*110.51* **	*35.62* **
1 -	Age	0.678 **	0.364 **	0.242 **	0.008
		(5.70)	(2.82)	(2.96)	(0.10)
2 -	Years of schooling	1.345 **	0.651	-0.235	-0.621 **
		(3.07)	(1.20)	-(0.68)	-(1.87)
3 -	Schooling x Age (x 10^{-2})	-4.041 **	-2.374 *	-0.063	1.026
		-(3.57)	-(1.73)	-(0.08)	(1.41)
4 -	Rural migrant	11.184 **	2.849	-1.805	-6.422 **
		(3.43)	(0.66)	-(0.53)	-(2.26)
5 -	Rural migrant x Age	-0.225 **	-0.042	0.073	0.135 **
		-(2.90)	-(0.41)	(0.85)	(1.98)
	Household Assets Variables	*5.27*	*4.44*	*0.73*	*1.24*
6 -	Non-Labor Income (x 10^{-2})	12.393	9.078	5.540	-0.088
		(1.16)	(0.69)	(0.64)	-(0.01)
7 -	Nonproductive assets	-0.071	-0.360 **	0.043	-0.041
		-(0.38)	-(1.90)	(0.25)	-(0.25)
8 -	Home business	2.340 **	0.809	0.495	0.947
		(2.01)	(0.50)	(0.51)	(1.09)
	Infrastructure Variables	*6.97* **	*2.24*	*7.76* **	*5.03* *
9 -	Adequate ceiling	-2.935 **	-1.574	-2.978 **	-2.018 **
		-(2.49)	-(1.06)	-(2.75)	-(2.14)
10 -	Adequate floors	-0.440	-1.287	0.543	-0.379
		-(0.37)	-(0.89)	(0.45)	-(0.40)
	Community Variables	*1.22*	*1.83*	*1.65*	*1.07*
11 -	Distance time to health service	1.619	3.630	-2.810	1.268
		(0.46)	(1.31)	-(1.22)	(0.82)
12 -	Waiting time to medical attention	1.216	-0.627	-0.362	-0.630
		(0.98)	-(0.46)	-(0.33)	-(0.66)
	Price Variables	*4.32*	*4.67*	*13.35* **	*1.81*
13 -	Potato price	6.410 **	7.609 *	4.532	0.995
		(2.07)	(1.82)	(1.23)	(0.41)
14 -	Milk price	3.014	-5.555	13.083 **	4.450
		(0.70)	-(0.85)	(3.51)	(1.33)

(continuation of table A.1)		Male		Females	
	Variables	illness	disability	illness	disability
	Regional Controls	*35.79* **	*3.65*	*66.99* **	*10.91* **
15 -	Costa	9.066 **	-0.241	11.790 **	4.865
		(3.91)	-(0.10)	(5.90)	(2.98) **
16 -	Other sierra	7.820 **	0.870	9.324 **	3.038 **
		(3.57)	(0.41)	(4.79)	(2.14)
17 -	Sierra sur	13.002 **	2.457	15.325 **	3.591 *
		(5.92)	(0.80)	(7.58)	(1.83)
18 -	Selva alta	5.024 **	3.187	10.216 **	3.455 *
		(2.09)	(1.12)	(3.97)	(1.80)
19 -	Selva baja	5.899 **	-1.469	8.728 **	1.225
		(2.04)	-(0.42)	(3.11)	(0.49)
20 -	May	26.582 **	19.412	13.337 **	6.974
		(1.84)	(1.57) *	(7.76)	(1.03)
21 -	June	3.874	-1.117	3.788 **	3.026 *
		(1.56)	-(0.59)	(2.13)	(1.59)
22 -	July	0.881	-5.739 **	1.947	2.011
		(0.35)	-(2.64)	(1.12)	(1.01)
23 -	Constant	-57.115 **	-27.053 **	-48.561 **	-24.068 **
		-(5.81)	-(2.12)	-(5.64)	-(3.15)
Log Likelihood		-3407.65	-1314.61	-4883.92	-2179.53
Global Chi-squared		274.23 **	89.70 **	284.36 **	79.35 **
Number of observations		3102	3102	3508	3508

(*) Statistically significant at 10% level of confidence.
(**) Statistically significant at 5% level of confidence.

Table A.2
Health Determinants for Peruvian urban males: Alternative dependent variables
(t-statistics in parentheses)

		Male			Female		
	Variables	Probit	Week-ill	Day-ill	Probit	Week-ill	Day-ill
	Individual Characteristics	*73.46* **	*87.29* **	*84.44* **	*117.71* **	*110.91* **	*110.51* **
1 -	Age	0.038 **	0.684 **	0.678 **	0.019	0.241	0.242 **
		(5.46)	(5.70)	(5.70)	(3.96)	(2.89)	(2.96)
2 -	Years of schooling	0.078 **	1.364 **	1.345 **	0.012	-0.250	-0.235
		(2.95)	(3.08)	(3.07)	(0.60)	-(0.71)	-(0.68)
3 -	Schooling x Age (x 10^{-2})	-0.220 **	-4.059 **	-4.041 **	-0.055	-0.038	-0.063
		-(3.36)	-(3.55)	-(3.57)	-(1.22)	-(0.05)	-(0.08)
4 -	Rural migrant	0.681 **	11.624 **	11.184 **	-0.144	-1.660	-1.805
		(3.24)	(3.57)	(3.43)	-(0.66)	-(0.48)	-(0.53)
5 -	Rural migrant x Age	-0.015 **	-0.234 **	-0.225 **	0.005	0.069	0.073
		-(2.89)	-(2.99)	-(2.90)	(0.92)	(0.79)	(0.85)
	Household Assets Variables	*5.81*	*5.56*	*5.27*	*1.49*	*0.56*	*0.73*
6 -	Non-Labor Income (x 10^{-2})	0.870	12.788	12.393	0.230	5.285	5.540
		(1.26)	(1.18)	(1.16)	(0.42)	(0.60)	(0.64)
7 -	Nonproductive assets	-0.004	-0.069	-0.071	0.000	0.020	0.043
		-(0.41)	-(0.38)	-(0.38)	(0.02)	(0.12)	(0.25)
8 -	Home business	0.143 **	2.410 **	2.340 **	0.070	0.430	0.495
		(2.20)	(2.07)	(2.01)	(1.20)	(0.43)	(0.51)
	Infrastructure Variables	*6.54* **	*6.86* **	*6.97* **	*7.7*	*7.19*	*7.76* **
9 -	Adequate ceiling	-0.159 **	-2.934 **	-2.935 **	-0.175	-2.894	-2.978 **
		-(2.29)	-(2.48)	-(2.49)	-(2.75)	-(2.66)	-(2.75)
10 -	Adequate floors	-0.044	-0.412	-0.440	0.009	0.501	0.543
		-(0.56)	-(0.35)	-(0.37)	(0.12)	(0.42)	(0.45)
	Community Variables	*1.55*	*1.28*	*1.22*	*3.89*	*1.68*	*1.65*
11 -	Distance time to health service	0.037	1.708	1.619	-0.297	-2.766	-2.810
		(0.17)	(0.48)	(0.46)	-(1.97)	-(1.20)	-(1.22)
12 -	Waiting time to medical attention	0.090	1.244	1.216	0.003	-0.450	-0.362
		(1.23)	(1.00)	(0.98)	(0.05)	-(0.41)	-(0.33)
	Price Variables	*6.96*	*4.35*	*4.32*	*14.21*	*13.36*	*13.35* **
13 -	Potato price	0.463 **	6.568 **	6.410 **	0.257	4.831	4.532
		(2.59)	(2.09)	(2.07)	(1.07)	(1.33)	(1.23)
14 -	Milk price	0.256	2.745	3.014	0.853	12.909	13.083 **
		(0.93)	(0.62)	(0.70)	(3.72)	(3.49)	(3.51)
	Regional Controls	*60.26* **	*36.23* **	*35.79* **	*56.64*	*67.12*	*66.99* **
15 -	Costa	0.564 **	9.154 **	9.066 **	0.653	11.756	11.790 **
		(4.21)	(3.94)	(3.91)	(5.20)	(5.91)	(5.90)
16 -	Other sierra	0.529 **	7.893 **	7.820 **	0.505	9.414	9.324 **
		(3.94)	(3.55)	(3.57)	(4.15)	(4.86)	(4.79)
17 -	Sierra sur	0.888 **	13.236 **	13.002 **	0.851	15.514	15.325 **
		(7.57)	(5.95)	(5.92)	(7.07)	(7.67)	(7.58)
18 -	Selva alta	0.281 **	5.167 **	5.024 **	0.511	10.225	10.216 **
		(2.13)	(2.09)	(2.09)	(2.65)	(3.96)	(3.97)
19 -	Selva baja	0.334 **	6.012 **	5.899 **	0.512	8.736	8.728 **
		(2.08)	(2.04)	(2.04)	(3.15)	(3.07)	(3.11)
20 -	May	0.468	25.423 *	26.582 **	0.960	13.401	13.337 **
		(0.89)	(1.68)	(1.84)	(5.71)	(7.38)	(7.76)
21 -	June	0.265 *	3.731	3.874	0.234	3.979	3.788 **
		(1.83)	(1.48)	(1.56)	(2.09)	(2.22)	(2.13)

(continuation of table A.2)

Variables	Male Probit		Week-ill		Day-ill		Female Probit		Week-ill		Day-ill	
22 - July	0.123		0.735		0.881		0.097		2.099		1.947	
	(0.86)		(0.29)		(0.35)		(0.88)		(1.20)		(1.12)	
23 - Constant	-3.599	**	-57.469	**	-57.115	**	-3.143		-48.431		-48.561	**
	-(6.28)		-(5.75)		-(5.81)		-(6.13)		-(5.65)		-(5.64)	
Log Likelihood	-1423.12		-2388.69		-3407.65		-1912.18		-3371.23		-4883.92	
Global Chi-squared	273.52	**	273.40	**	274.23	**	369.43		283.91		284.36	**
Number of observations	3094		3102		3102		3493		3508		3508	

(*) Statistically significant at 10% level of confidence.
(**) Statistically significant at 5% level of confidence.

Table A.3
Health Determinants for urban adults by gender: Life-cycle effects
(t-statistics in parentheses)

		Male				Females	
	Variables	Full Sample	Young	Mature	Old	Full Sample	You
	Individual Characteristics	*84.44* **	*16.65* **	*7.64*	*7.28*	*110.51* **	
1 -	Age	0.678 **	1.485 **	0.815	0.083	0.242 **	-
		(5.70)	(2.15)	(1.40)	(0.16)	(2.96)	-(
2 -	Years of schooling	1.345 **	3.303 **	1.660	-1.310	-0.235	-
		(3.07)	(2.18)	(0.92)	-(0.43)	-(0.68)	-(
3 -	Schooling x Age (x 10^{-2})	-4.041 **	-0.129 **	-0.047	0.011	-0.063	
		-(3.57)	-(2.05)	-(1.00)	(0.19)	-(0.08)	(
4 -	Rural migrant	11.184 **	11.919	9.208	-9.848	-1.805	-
		(3.43)	(1.20)	(0.64)	-(0.39)	-(0.53)	-(
5 -	Rural migrant x Age	-0.225 **	-0.262	-0.180	0.148	0.073	
		-(2.90)	-(0.66)	-(0.48)	(0.32)	(0.85)	(
	Household Assets Variables	*5.27*	*8.76*	*2.60*	*5.52*	*0.73*	
6 -	Non-Labor Income (x 10^{-2})	12.393	0.344	-0.275	0.081	5.540	
		(1.16)	(2.44)	-(1.42)	(0.31)	(0.64)	(
7 -	Nonproductive assets	-0.071	0.137	-0.100	-0.075	0.043	-
		-(0.38)	(0.60)	-(0.36)	-(0.18)	(0.25)	-(
8 -	Home business	2.340 **	1.831	1.325	5.982 **	0.495	
		(2.01)	(1.15)	(0.79)	(2.29)	(0.51)	(
	Infrastructure Variables	*6.97* **	*0.51*	*3.98*	*5.75* **	*7.76* **	
9 -	Adequate ceiling	-2.935 **	-1.014	-3.282 *	-6.897 **	-2.978 **	-
		-(2.49)	-(0.64)	-(1.68)	-(2.39)	-(2.75)	-(
10 -	Adequate floors	-0.440	-0.240	-0.963	0.400	0.543	
		-(0.37)	-(0.15)	-(0.43)	(0.14)	(0.45)	(
	Community Variables	*1.22*	*1.83*	*1.88*	*1.23*	*1.65*	
11 -	Distance time to health service	1.619	-2.034	4.681	7.015	-2.810	-
		(0.46)	-(0.89)	(0.77)	(1.09)	-(1.22)	-(
12 -	Waiting time to medical attention	1.216	1.548	1.887	0.389	-0.362	-
		(0.98)	(1.16)	(1.21)	(0.18)	-(0.33)	-(

(continuation of table A.3)			Male			
Variables	Full Sample	Young	Mature	Old	Full Sample	You
Price Variables	*4.32*	*0.24*	*5.33*	*8.91*	*13.35* **	
13 - Potato price	6.410 **	1.923	9.139 **	14.318 **	4.532	
	(2.07)	(0.48)	(2.19)	(2.32)	(1.23)	(
14 - Milk price	3.014	1.981	-1.554	12.421 **	13.083 **	1
	(0.70)	(0.29)	-(0.31)	(2.01)	(3.51)	(
Regional Controls	*35.79* **	*23.19* **	*21.15* **	*19.54* **	*66.99* **	
15 - Coast	9.066 **	8.364 **	6.394 **	14.196 **	11.790 **	1
	(3.91)	(3.11)	(2.01)	(3.87)	(5.90)	(
16 - Other Sierra	7.820 **	6.902 **	6.796 **	10.052 **	9.324 **	
	(3.57)	(3.03)	(2.22)	(2.40)	(4.79)	(
17 - Sierra sur	13.002 **	11.828 **	11.942 **	16.769 **	15.325 **	1
	(5.92)	(4.44)	(4.02)	(3.91)	(7.58)	(
18 - Selva alta	5.024 **	4.764 *	3.466	6.431	10.216 **	
	(2.09)	(1.83)	(1.04)	(1.52)	(3.97)	(
19 - Selva baja	5.899 **	8.651 **	2.977	2.174	8.728 **	
	(2.04)	(2.33)	(0.82)	(0.45)	(3.11)	(
20 - May	26.582 **	-70.246 **	132.264 **	13.845 **	13.337 **	1
	(1.84)	-(8.84)	(12.99)	(2.05)	(7.76)	(
21 - June	3.874	5.933 **	8.063 **	-5.765	3.788 **	
	(1.56)	(2.24)	(3.05)	-(1.09)	(2.13)	(
22 - July	0.881	3.712	5.392 **	-10.273 *	1.947	
	(0.35)	(1.35)	(2.44)	-(1.83)	(1.12)	(
23 - Constant	-57.115 **	-71.636 **	-62.271 **	-42.547	-48.561 **	-4
	-(5.81)	-(3.42)	-(2.37)	-(1.45)	-(5.64)	-(
Log Likelihood	-3407.65	-1404.98	-1152.85	-822.81	-4883.92	-20
Global Chi-squared	274.23 **	86.94 **	84.42 **	96.44 **	284.36 **	
Number of observations	3102	1548	963	591	3508	

(*) Statistically significant at 10% level of confidence.
(**) Statistically significant at 5% level of confidence.

Appendix B

Table B.1: Descriptive Statistics of the Peruvian Urban Population

	Total Population			Male Population			
	Total	Male	Female	Unemployed	Wage Earners	Self Employed	Unemploy
Sample size	6610	3102	3508				
Labor force participation (%)	40.3	45.1	22.6	23.3	45.1	31.6	5
Dependent Health Variables							
Days sick	2.322	1.885	2.709	1.752	1.590	2.403	2.
	(0.072)	(0.096)	(0.106)	(0.203)	(0.126)	(0.190)	(0.1
Hourly Income Variables /a							
Last 12 months					3.157	4.147	
					(0.113)	(0.257)	
Last 7 days					3.111	4.190	
					(0.109)	(0.271)	
Individual Variables							
Age (in years)	32.818	32.800	32.834	23.830	33.676	38.171	31.
	(0.153)	(0.225)	(0.208)	(0.413)	(0.302)	(0.372)	(0.3
Tenure (in years)	4.345	5.901	2.969	1.028	6.723	8.324	1.
	(0.087)	(0.145)	(0.098)	(0.119)	(0.208)	(0.297)	(0.0
Years of schooling	9.920	10.401	9.495	10.436	10.874	9.700	9.
	(0.050)	(0.067)	(0.073)	(0.115)	(0.102)	(0.128)	(0.0
Primary education dummy	0.196	0.169	0.221	0.086	0.151	0.255	0.
	(0.005)	(0.007)	(0.007)	(0.010)	(0.010)	(0.014)	(0.0
Secondary education dummy	0.489	0.517	0.465	0.604	0.496	0.482	0.
	(0.006)	(0.009)	(0.008)	(0.018)	(0.013)	(0.016)	(0.0
High education dummy	0.275	0.297	0.255	0.289	0.340	0.242	0.
	(0.005)	(0.008)	(0.007)	(0.017)	(0.013)	(0.014)	(0.0
Rural migrant	0.213	0.211	0.214	0.080	0.221	0.300	0.
	(0.005)	(0.007)	(0.007)	(0.010)	(0.011)	(0.015)	(0.0
Public worker dummy	0.092	0.113	0.074	0.007	0.239	0.010	
	(0.004)	(0.006)	(0.004)	(0.003)	(0.011)	(0.003)	
Blue collar worker dummy	0.159	0.255	0.075	0.032	0.513	0.050	
	(0.005)	(0.008)	(0.004)	(0.007)	(0.013)	(0.007)	

	Total Population			Male Population			
	Total	Male	Female	Unemployed	Wage Earners	Self Employed	Unemploy
Firm Size Variables							
Micro scale firm dummy	0.119	0.121	0.115		0.156	0.046	
	(0.323)	(0.327)	(0.319)		(0.363)	(0.210)	
Small scale firm dummy	0.087	0.083	0.094		0.127	0.015	
	(0.282)	(0.275)	(0.291)		(0.333)	(0.123)	
Medium scale firm dummy	0.180	0.207	0.138		0.327	0.027	
	(0.384)	(0.406)	(0.345)		(0.469)	(0.161)	
Large scale firm dummy	0.047	0.052	0.039		0.083	0.008	
	(0.211)	(0.222)	(0.193)		(0.276)	(0.090)	
Income Variables							
Rents + Transfers /a	29.847	24.176	34.862	37.982	20.741	18.891	37.
	(0.848)	(1.118)	(1.249)	(2.804)	(1.403)	(2.013)	(1.7
Non productive stock /a	32095	32366	31856	50611	24879	29597	33
	(1645)	(2512)	(2161)	(7846)	(2053)	(4566)	(30
Household business /a	0.724	0.729	0.720	0.679	0.564	1.000	0.
	(0.005)	(0.008)	(0.008)	(0.017)	(0.013)	(0.000)	(0.0
Community Variables							
Unemployment rate (by district)	0.925	0.925	0.926	0.917	0.931	0.921	0.
	(0.001)	(0.002)	(0.001)	(0.004)	(0.002)	(0.003)	(0.0
Hired rate (by province)	0.496	0.497	0.496	0.493	0.529	0.453	0.
	(0.002)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.0
Hired rate (by district)	0.447	0.447	0.446	0.440	0.458	0.435	0.
	(0.001)	(0.002)	(0.002)	(0.005)	(0.003)	(0.004)	(0.0
Housing Infrastructure Variables							
Adequate ceiling	0.538	0.540	0.537	0.636	0.522	0.493	0.
	(0.006)	(0.009)	(0.008)	(0.018)	(0.013)	(0.016)	(0.0
Adequate floor	0.771	0.770	0.771	0.819	0.766	0.741	0.
	(0.005)	(0.008)	(0.007)	(0.014)	(0.011)	(0.014)	(0.0
Health Price Variables							
Distance time to health center /b	0.364	0.367	0.360	0.387	0.356	0.371	0.
	(0.002)	(0.003)	(0.003)	(0.008)	(0.005)	(0.006)	(0.0
Waiting time to medical attention /b	1.056	1.051	1.060	1.037	1.046	1.074	1.
	(0.006)	(0.008)	(0.008)	(0.016)	(0.013)	(0.014)	(0.0

	Total Population			Male Population			
	Total	Male	Female	Unemployed	Wage Earners	Self Employed	Unemploy
Price Variables							
Potato price /a	0.862	0.866	0.859	0.842	0.862	0.890	0.
	(0.003)	(0.004)	(0.004)	(0.008)	(0.006)	(0.008)	(0.(
Milk price /a	1.489	1.490	1.488	1.486	1.497	1.484	1.
	(0.002)	(0.003)	(0.003)	(0.006)	(0.005)	(0.005)	(0.(
Regional and Date Variables							
Costa	0.228	0.217	0.237	0.178	0.233	0.224	0.
	(0.005)	(0.007)	(0.007)	(0.014)	(0.011)	(0.013)	(0.(
Other sierra	0.125	0.120	0.128	0.170	0.104	0.106	0.
	(0.004)	(0.006)	(0.006)	(0.014)	(0.008)	(0.010)	(0.(
Sierra sur	0.101	0.105	0.097	0.118	0.094	0.112	0.
	(0.004)	(0.006)	(0.005)	(0.012)	(0.008)	(0.010)	(0.(
Selva alta	0.059	0.060	0.059	0.058	0.056	0.067	0.
	(0.003)	(0.004)	(0.004)	(0.009)	(0.006)	(0.008)	(0.(
Selva baja	0.105	0.107	0.103	0.082	0.095	0.142	0.
	(0.004)	(0.006)	(0.005)	(0.010)	(0.008)	(0.011)	(0.(
May	0.001	0.001	0.001	0.003	0.001	0.000	0.
	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)	(0.000)	(0.(
June	0.509	0.503	0.515	0.512	0.487	0.518	0.
	(0.006)	(0.009)	(0.008)	(0.019)	(0.013)	(0.016)	(0.(
July	0.443	0.451	0.436	0.452	0.459	0.438	0.
	(0.006)	(0.009)	(0.008)	(0.019)	(0.013)	(0.016)	(0.(

* The numbers indicated are the mean values of the variables
** The numbers in parentheses are the standard errors of the variables
/a Nuevos soles of June 94
/b In hours